View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by OpenCommons at University of Connecticut

UCONN LIBRARY

University of Connecticut
OpenCommons@UConn

Master's Theses | University of Connecticut Graduate School

12-16-2016

The Inter- and Intra-Rater Reliability of Athletic Trainers at Identifying ACL Injury Risk Factors Using the Cutting Screening Tool

Kelly Buckholz Smallbeck
University of Connecticut - Storrs, kelly.buckholz@gmail.com

Recommended Citation

Buckholz Smallbeck, Kelly, "The Inter- and Intra-Rater Reliability of Athletic Trainers at Identifying ACL Injury Risk Factors Using the Cutting Screening Tool" (2016). *Master's Theses*. 1038.
https://opencommons.uconn.edu/gs_theses/1038

This work is brought to you for free and open access by the University of Connecticut Graduate School at OpenCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of OpenCommons@UConn. For more information, please contact opencommons@uconn.edu.

The Inter- and Intra-Rater Reliability of Athletic Trainers at Identifying ACL Injury Risk Factors Using the Cutting Screening Tool

Kelly Michelle Buckholz Smallbeck

B.S. in Exercise and Sport Science, Oregon State University, 2010

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

COPYRIGHT

© 2016

Kelly Michelle Buckholz Smallbeck

ALL RIGHTS RESERVED

APPROVAL PAGE

Master of Science Thesis

The Inter- and Intra-Rater Reliability of Athletic Trainers at Identifying ACL Injury Risk Factors Using the Cutting Screening Tool

Presented By

Kelly Michelle Buckholz Smallbeck, B.S.

Major Advisor____
Lindsay J. DiStefano

Associate Advisor____
Douglas J. Casa

Associate Advisor____
Thomas Trojian

Associate Advisor____
Carl M. Maresh

University of Connecticut

2016

ACKNOWLEDGEMENTS

Dr. DiStefano – Words cannot describe how grateful I am for your guidance and support. I am extremely blessed to have an adviser like you, who always pushed me to continue to work and fight. You believed in me when I had almost stopped believing in myself, and that support has led me to accomplish things I never thought I could.

Dr. Casa – Thank you so much for taking a chance on me and welcoming me into such a wonderful program at the University of Connecticut. You made it easy to believe that moving across the country would be a wonderful chance, and I wouldn't trade my time here for anything.

Dr. Maresh – Thank you for being a part of my thesis committee. Your dedication to furthering the department and the university is truly inspiring. Thank you for helping me believe that a West Coast transplant could succeed in New England.

Dr. Trojian – Thank you for being a part of my thesis committee. Your standards for educating others pushed me to want to be better. I appreciate all the work you have done for this department and for your time and dedication to my thesis.

This would not have been possible without the patient (and sometimes impatient) support of my family and friends. Your love and support has guided me through the last two years. And to Scott Smallbeck, you are more than I asked for and just what I need. Thank you for always knowing what I'm thinking, both good and bad, and handling it anyway.

TABLE OF CONTENTS

TABLES AND FIGURES

Tables

Figures

ABSTRACT

Background: ACL injuries are a frequently talked about injury in sports and have a multitude of negative side effects. A majority of injuries occur during single-leg cutting motions. Clinical screening tools are an effective way to identify those demonstrating more ACL injury risk factors. Recent research has created a variety of clinical screening tools to identify ACL injury risk factors, however none screen athletes during a single-leg cutting motion.

Hypothesis: A clinical screening tool using a single-leg cutting motion (the Cutting Screening Tool, or CST) will reliably identify ACL injury risk factors when used by athletic trainers.

Methods: Six currently certified, licensed, and practicing athletic trainers (4 Female, 2 Male; 3 "Novice" with less than five years experience, 3 "Veteran" with more than 10 years experience) in the state of Connecticut were recruited to rate 20 subjects performing three trials of a single-leg cutting task from the frontal and sagittal planes. This data was used to determine inter-rater reliability. After a minimum waiting period of three days, the raters re-scored five subjects a second time to assess intra-rater reliability.

Results: The CST had overall poor inter-rater reliability ($ICC_{(2,1)} = 0.32$, SEM = 1.42). Inter-rater reliability of specific items on the CST varied from poor to excellent. One system of data reduction (Cohen's kappa statistic) occasionally had low inter-rater reliability because of high percentage of chance agreement. Intra-rater reliability was moderate when comparing Novice vs. Veteran raters ($ICC_{(2,1)} = 0.64$, SEM = 0.69) Novice raters had excellent intra-rater reliability ($ICC_{(2,1)} = 0.99$, SEM = .10) while Veteran raters had poor intra-rater reliability ($ICC_{(2,1)} = 0.38$, SEM = 1.72). One Veteran rater was removed from certain calculations due to discrepancies in re-scoring subjects.

Conclusions: The CST is not currently ready for use as a clinical screening tool, however it will be a valuable method of identifying those at greater risk for ACL injury following further research.

Keywords: ACL injury; clinical screening tool; single-leg; cutting motion; injury risk

REVIEW OF LITERATURE

Epidemiology

Official rates for anterior cruciate ligament (ACL) injuries vary greatly, both internationally and nationally. Over a five-year period, 80% of knee surgeries in New Zealand were reported as having some type of ACL component.[1] De Loës *et al.*[2] found in a 2000 study Swiss youth (ages 14-20) had an injury rate of one anterior/posterior cruciate ligament (PCL) tear for every 5,000 participants. In a National Health Care Survey in 2003, Marshall, *et al.*[3] discovered that approximately 1 in 90 visits to a physician, hospital emergency room, or hospital-based outpatient center in the United States was in some way related to a cruciate ligament injury. The Marshall data does not separate ACL from PCL.[3] There is no national database currently tracking ACL injuries as distinct from PCL injuries in the United States, unlike the data compiled in New Zealand. Many countries, including the United States, currently lack an accurate national injury data tracking system for the general population, including cruciate ligament injuries. In addition, the U.S. data does not count the number of initial visits to a medical facility--merely those that occurred during the time of the survey. Cruciate injuries are a long-term injury, often requiring multiple doctor or hospital visits. There is currently no reliable data on the number of initial visits for an injury, so we must estimate. Based on this information, Marshall, *et al.* [3] estimated 200,000 U.S. cruciate ligament injuries per year. This new data more than triples older data previously reported.[3]

The National Collegiate Athletic Association (NCAA) has been compiling injury data on collegiate athletes in all divisions (I, II, and III) since 1982. Over the course of 16 years (fall of 1988-spring 2004) ACL injuries were the third most common injury, behind ankle sprains and concussions.[4] During those years, football and women's basketball reported the most injuries; all 15 sports recorded nearly 5,000 ACL injuries.[4] Those 16 years showed an average of 313 injuries per year, with a significant 1.3% increase yearly.[4] The sports where ACL injuries composed the highest percentage of injuries for the sport were women's basketball, gymnastics and lacrosse (4.9, 4.9 and 4.3%, respectively).[4] However, some of the highest incidences occur in women's basketball and soccer; for men, football.[4] Men's football has the highest occurrence of ACL injuries across all sports, with 2,538 total injuries for both fall and spring football.[4] But the percentage of ACL injuries compared to total injuries is much lower than other sports, only 3%.[4] This is significant because while other teams may have 10-25 athletes on the roster, football rosters often range from 50-100 players per team. With all of these injuries occurring, it is important to recognize how it affects the athletes long-term.

The Costs and Consequences of Injury

There are many costs and consequences associated with ACL injuries. Three of the most impactful for athletes are financial cost, reduction in physical capabilities both short and long-term, and early onset of osteoarthritis. De Loës *et al.*[2], found that ACL/PCL injuries incur the highest financial cost of all knee injuries in Sweden. The cost reported by De Loës *et al.*[2] was significantly lower than the national average cost of ACL repairs reported by other countries, as well.[5] In 2003, Belgium spent almost 1,400 euros (approximately US$2,000 then) on every ACL repair, which was low compared to national averages around the globe.[5-7] Even with the lower cost, it was still the most expensive sport-related injury in the country.[6] Gianotti, *et al.*[1] reported that the average cost of ACL reconstruction surgery alone in New Zealand as of 2009 was around US$17,000 per surgery.[1] This does not include rehabilitation and is on par with averages in the United States.[3, 5]

Silvers, *et al.*[5] estimates the national average for ACL reconstruction in the United States to be around two billion dollars annually. The national average for ACL reconstruction and rehabilitation in the United States is an estimated three billion annually.[7] Combined with the data from Marshall, *et al.*[3], it is estimated that the average cost of ACL surgery in the United States is approximately $15,000 per reconstruction. Cumps, *et al.*[8] estimated that the average cost of the ACL graft alone cost $5,000. Costs and consequences of ACL injury are not limited to the financial realm, but also the physical.

Athletes who suffer an ACL injury and undergo reconstruction are typically removed from competition for a minimum of five to six months depending on the sport.[9] Manske, *et al.*[9] reviewed the most current literature regarding current best practices for the rehabilitation of athletes post-ACL reconstruction and return to play protocols. They highlighted the controversy surrounding the standard six-month return to play protocol when there is little evidence supporting this timeline.[9] Research shows that there are often decrements in strength, proprioception, postural stability, and bilateral limb balance not specific to the injured knee for two years post-reconstruction and beyond.[10-12] As a result of this research, clinicians are now encouraging a longer rehabilitation timeline to allow for full recovery of not just the knee but the athlete post-ACL reconstruction.[10-12] Some athletes experience difficulty returning to the prior level of competition post-ACL injury and reconstruction.[10-15] The aftereffects of an ACL injury continue into the patient's future.

Studies show that people who sustain an ACL injury have an increased chance of developing knee osteoarthritis (OA) later in life.[6, 16-18] Lohmander, *et al.*[17] studied 103 females 12 years after an ACL repair, using a weight-bearing radiograph and questionnaire to determine presence of OA. Those who participated had a mean age of 31. Of 67 radiographs, 55 showed visible changes in their knee, 34 of which qualified as having OA: (50%).[17] Of the 84 subjects who completed questionnaires, 75% reported some type of knee-related problem that was affecting their quality of life.[17] Another study showed that immediately following ACL injury, the cartilage and collagen in the knee begins to have visible changes as a result of the rupture.[6] Nelson, *et al.*[6] compared 50 post-ACL-repair cartilage samples to 21 control cartilage samples. Over half were one year or more post-surgery. Even in samples less than one year old, there were significant changes in the collagen of the knee cartilage, which slowly decreased over time. Unfortunately, as the changes in the cartilage decreased, the Mankin score (for cartilage degeneration overall) increased.[6]

In another study (Kessler, *et al.*[18]), 136 patients with an isolated ACL rupture were evaluated in clinical, radiological and knee grading scales that are internationally accepted (Tegner, IKDC, Kellgren and Lawrence). Of those contacted, 27 had undergone a second repair and were excluded; 60 had their knees reconstructed; and 49 were treated non-operatively. Of those with reconstructions, increased stability was reported; however, their rate of OA was higher (42%) compared to the non-operative group (25%).[18] Overall, 24% of the 136 participants developed OA.

Brown, *et al.*[16] examined the nationwide prevalence and cost associated with OA when compared to the general population. They predict that 12% of the population suffers from OA as a result of a posttraumatic event in the lower extremity in previous years.[16] Comparing the relationship between rheumatoid arthritis and the costs associated with each type of illness, they calculated that posttraumatic OA has an annual cost of $3.06 billion in the U.S. alone.[16] This study was not limited to the knee. Though ACL injuries are not specific to one gender or another, females are more susceptible to ACL injury and therefore the associated problems.

Gender Differences

Following the passage of Title IX legislation in 1972, there was a dramatic increase in female participation in sports at all levels.[19] Lopiano found that in 1972, at the time of Title IX's passage, there were over 800,000 females participating in sports at the high school level and almost 33,000 at the college level. By 1996 these numbers had increased to 2.4 million females participating at the high school level and 128,000 at the college level.[19] The same dramatic increase in participation in sports is not seen in males, whose high school and college participation numbers have remained steady.[19] With the increase in female participation, there was an understandable increase in the frequency of female injuries.[20] However, this increase in frequency did not translate into an increased rate of all injuries.[20] Unfortunately, data shows that females are more predisposed to ACL injuries than males.[21, 22]

Female athletes are 3.5 times more likely to tear their ACLs than their male counterparts.[21] Of the four college sports with the highest rate of ACL injury, three of them are female sports.[4] The teams include, in order of injury rate, Women's gymnastics, Men's spring football, Women's soccer and Women's basketball. Women's gymnastics and Men's spring football had the highest rates, at .33 per 1,000 exposures.[4] Women's soccer has an injury rate of .28 per 1,000, followed by basketball with an ACL injury rate of .23 per 1,000 exposures.[4] After the top four sports, there is a marked drop off in ACL injury rate.[4] The sport with the fifth highest rate is Men's football, with a rate of .18 per 1,000 exposures.[4]

The gender discrepancy is not limited to collegiate-age athletes. In studies of both Texas and New Jersey high school athletes, females were at least 3.5 times more likely to sustain knee injuries per hour of injury risk exposure.[21] Male rate of sport participation is currently higher than that of females.[19] In spite of the fact that males typically receive more injuries overall, the rate of female injury is much higher because the ratio of female injuries to female participants is greater than the ratio of male injuries to male participants. [2, 3, 21] (ACL injuries or all injuries?) These statistics do not carry across all ages, but are affected most strongly with puberty.

The incidence of ACL injury in female athletes increases dramatically following the changes associated with puberty when compared to males, with the highest numbers reported between the ages of 16-18.[3, 23-25] The gender differences continue into adulthood. Agel, *et al.*[22] found that female collegiate basketball players were 4.6 times more likely to tear their ACLs than male collegiate basketball players using data collected over 13 years. This difference is also present in soccer players, where female athletes were 2.78 times more likely to suffer an ACL injury than males.[22]

The high number of negative consequences and costs associated with ACL injury necessitate reliably identifying risk factors causing an increase in ACL injury. In order to do this, how risk factors increase chance of ACL injury must be understood. Risk factors typically divide into four specifications: anatomical design, hormonal influences, specific body and knee positions, and proprioceptive reactions present during a mechanism of injury that contribute to ACL overload. Mechanism of Injury (MOI) is the movement an athlete is performing when an ACL injury occurs. General sport-associated motions are not problematic unless they combine with present risk factors to cause an injury, becoming a MOI. Types of MOI include deceleration, landing, cutting, jumping, or other direction changes.[26] The presence of risk factors during sport-associated motions such as jumping and cutting does not guarantee an injury will occur. However, risk factors can build upon each other in a domino effect to overload the ACL, causing injury.

Risk factors are categorized into either non-modifiable or modifiable. Non-modifiable risk factors are contributors to ACL injury that modern science is currently unable to change, either because of the way our bodies are built or because of the way systems in the body affect them. Types of non-modifiable risk factors include genetic, anatomical, and hormonal risk factors. Modifiable risk factors are contributors to ACL injury that can be changed and acted upon. Currently the modifiable risk factors are biomechanical, or how the body moves and reacts. These include body position at the time of injury and proprioceptive reactions. Proprioceptive reactions are the timing, reactive ability, and ability of the body to react to outside stimuli. By understanding these risk factors, it allows us to better identify and correct them, minimizing not only the chance of ACL injury, but the chance of repeat injury as well.

Risk Factors

Anatomical Risk Factors

Anatomical risk factors are those influenced by the anatomical structure of the body itself, rather than hormonal or biomechanical systems. They cannot be modified, but must be understood to help determine risk. Early in ACL injury research, a main anatomical risk factor was believed to be the Quadriceps angle (Q-angle). The Q-angle is the angle of the center of the knee joint relative to the center of the hip joint. It was believed that an increased Q-angle caused by wider hips (as occurs primarily in women) increased one’s likelihood of injury. This has since been proven to be false.[27] What has shown to be a slightly more impacting factor in terms of bony anatomy is the ratio between the width of the pelvis and the length of the femur. An increased pelvic-width to femur length ratio has shown to be statistically significant in its relation to both static and dynamic stresses placed on the knee.[27] Another anatomical factor affecting ACL injury risk is the size of the femoral notch and the space relative to the size of the ACL. Women's femoral notches are not only smaller, but the size of a man's ACL is proportionate to the width and area of his femoral notch, whereas the ACL in a woman is more likely to be disproportionate, putting her at increased risk.[28]

In soft tissue anatomy, one risk factor that contributes to the risk of ACL injury is the fact that female ACLs are shorter, have a smaller cross-sectional area, and less volume than male ACLs.[28] This is true even when accounting for the height and weight difference between men and women. Female ACLs also have a lower elastic threshold compared with those of men, and a lower load level failure, making them more likely to rupture when compared to that of an ACL in a male.[29]

Hormonal Risk Factors

There is a current controversy involving ACL injuries and hormonal influence. One debate is whether or not the fluxuation of sex hormones during a woman's normal menstrual cycle causes changes in the laxity of the joints as well. It is thought that when a woman is menstruating (or premenstrual), the fluxuation in hormones loosen ligaments and make them more likely to tear.[30, 31] Slauterbeck, *et al.*[31] studied 38 female athletes (ranging in age from middle school to college) starting at the time of ACL injury, and measured their saliva for hormone levels for three years. The aim of the study was to determine whether or not estradiol and progesterone levels can indicate when in the menstrual cycle the injury occurred. Of the 27 who provided menstrual histories at the time of injury, 10 participants were injured either immediately before, or one to two days into, the beginning of their period.

Arendt, *et al.*[30] monitored 83 women for three years, dividing them into "on birth control" and "off birth control" groups. They had athletic trainers monitor the athletes for number of injuries and menstrual phase at time of injury. There was a significant correlation between the number of injuries and the phase in menstrual cycles:[30] most injuries occurred in the two weeks leading up to and beginning menstruation.[30, 30]

There was no difference between the numbers of injuries in the birth control group vs. non-birth control group. Hormones do play a role in female ACL injuries, making it more vital that correctable risk factors are addressed. Minimizing the effects of modifiable risk factors prevents a cumulative effect with non-modifiable risk factors to increase ACL injury risk.

Biomechanical Risk Factors

Many different biomechanical factors can influence ACL injury risk. Biomechanical risk factors are those that occur due to movement of the body, and are typically correctable. Boden, *et al.*[26] demonstrated that over half of ACL injuries that occur are non-contact. Non-contact means there is no direct contact between the knee and external forces; for example, there is no traumatic collision with another player that tears the ACL. The majority of these non-contact ACL injuries involve some type of running, cutting or landing mechanism of injury.[26] In a study analyzing videos of ACL injuries, 73% of them occurred in non-contact settings.[26]

When examining what stresses the ACL, in-vitro studies are the gold standard. In-vitro studies allow sensors to be placed on the ACL of a cadaver, giving quantitative force measurements when the ACL is loaded through different knee positions. Forces are applied through manipulation of the tibia and femur imitating specific musculature and dynamic motions. By controlling the knee joint in this method, researchers are able to examine the stresses placed on the ACL when specific motions are applied and combined, ultimately to failure of the ligament. In-vitro studies give us a better picture about what motions stress the ACL, including Anterior Tibial Shear Force (ATSF).

ATSF is a type of force that occurs around the knee; it is the anterior motion of the femoral condyles on the tibial plateau. The purpose of the ACL is to reduce ATSF. Without an ACL, little prevents the femur from trying to slide off the superior surface of the tibia. ATSF isn't harmful, and in fact is necessary to normal movement. However, too much ATSF can overload the ACL, causing a tear. ATSF is reduced by increasing knee flexion during motion.[32]

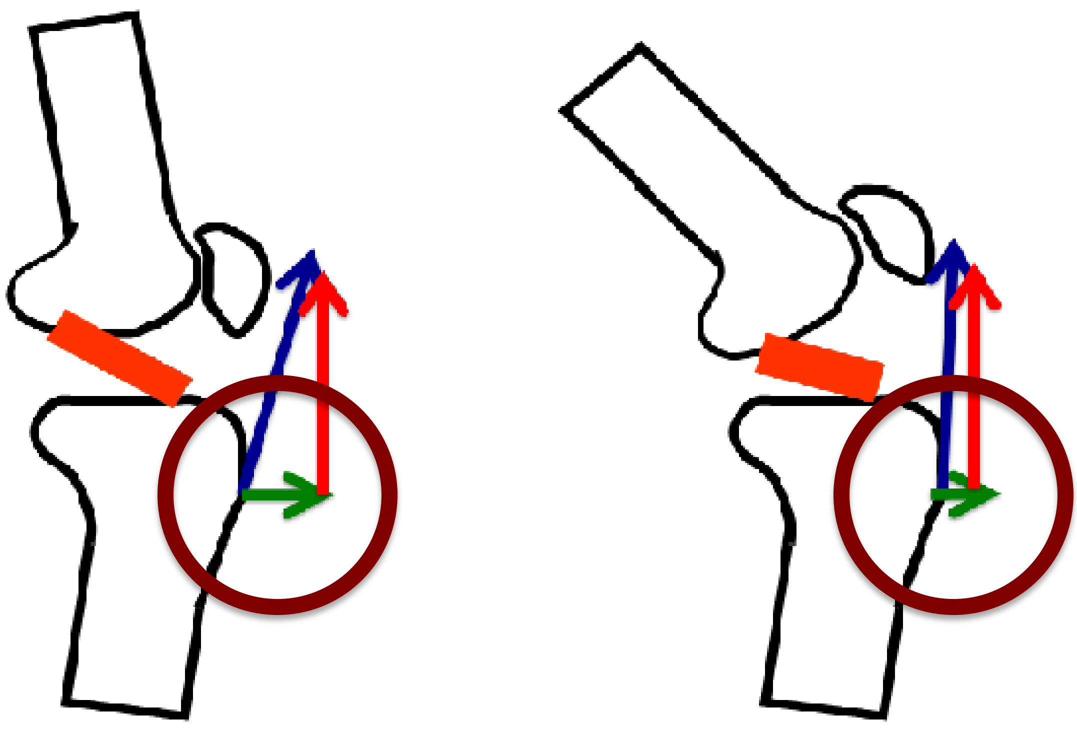

Figure 1. An example of the Anterior Tibial Shear Force (ATSF) on the ACL (left), which is reduced with increased knee flexion (right).

Markolf, *et al.*[32] found that at full extension the ACL is receiving 150% of the ATSF stresses when force is applied to it constantly compared to when the knee is flexed at approximately 30 degrees. Markolf demonstrated through an in-vitro study, that internal tibial rotation combined with ATSF, greatly increased the strain on the ACL when the knee was flexed under 20 degrees.[32] External tibial rotation combined with ATSF was increased with the knee was flexed less than 10 degrees.[32] The addition of valgus stress to ATSF also increased the stress on the ACL at any degree of knee flexion greater than five.[32] Currently there is a shift in thinking towards valgus stress around the knee and its role in ACL injury.

Previous research has used the term *knee valgus* (KV) to refer to the medial motion of the knee during a running, jumping, cutting or landing task. However, the more appropriate term that is now coming into use is *medial knee displacement* (MKD). True KV occurs primarily when the knee is extended rather than flexed, and is most easily seen when performing the Valgus Stress Test on the Medial Collateral Ligament (MCL). Medial knee displacement is an appropriate term because it is not limited to pure KV in a static position; rather it incorporates a combination of KV, hip internal rotation, tibial rotation, and knee flexion during a dynamic motion.[32] The updated terminology allows us a more broad term with which to identify ACL injury risks. In-vitro studies give us valuable knowledge on what stresses the ACL, but less information on how that is translated into dynamic movement in live subjects. Other types of research are needed to bridge the gap between in-vitro studies and ACL injury prevention.

To bridge the gap between ACL stresses and ACL prevention, researchers began examining videos of athletes at the time of ACL injury and the position of their bodies. In one study, Ireland[33] examined video footage from over the course of a decade and observed what she calls “The Position of No Return” during non-contact ACL injury.[33] Ireland observed that at the point of non-contact ACL injury, a majority of athletes demonstrated a forward flexed trunk rotated opposite the direction of a cut, an internally rotated hip, decreased knee flexion, increased MKD, and an externally rotated tibia.[33]

In a 2001 study of 20 ACL injuries in alpine skiing, Bere, *et al.*[34] examined footage of skiers tearing their ACLs during the course of three skiing seasons. Three main MOIs for ACL injury were identified.[34] Slip-catch injuries (n=10), involved a loss of balance during a turn, forcing the outside leg into hip internal rotation and MKD while trying to regain balance.[34] Dynamic snowplow injuries (n=3), were similar to slip-catch injuries, except during the turn, the inside leg was forced into internal rotation and MKD while attempting to regain balance.[34] The third category, landing back-weighted (n=4), involved landing on the back half of the skis, forcing the tibia and fibula anteriorly during initial contact of the skis with the ground (also known as deep knee flexion), tearing the ACLs.[34] Two of the three other ACL injuries occurred when the involved leg was forcefully externally rotated at the foot.[34] The final injury occurred during a crash and would qualify as a contact injury.[34]

Krosshaug, *et al.*[35] used video footage of athletes tearing their ACLs from multiple camera angles (basketball: four cameras, team handball: three cameras, and downhill skiing: one camera). These videos were matched with skeletal models frame by frame to the motion of the athlete. When researchers were satisfied the skeleton was as accurate as possible to the footage, they extrapolated kinematics and kinetics throughout the motion.[35] The basketball player was performing a single-leg 10 degree cut while receiving a pass at the time of injury.[35] Knee flexion at initial contact was 13 degrees with no MKD.[35] At 30 milliseconds (ms) post initial contact, knee flexion had increased to 35 degrees, with a rapid MKD to 14 degrees.[35] (30 ms: 22 degree increase in knee flexion, 14 degree increase in MKD). The team handball player was performing a single-leg right-to-left cutting maneuver of approximately 67 degrees at the time of injury.[35] At initial contact, the handball player had 26 degrees of hip abduction, 11 degrees of knee flexion, with no MKD.[35] When vertical force reached its peak 40 ms later, knee flexion had increased to 31 degrees and the athlete now had a MKD of 15 degrees.[35] (40 ms: 20 degree increase in knee flexion, 15 degree increase in MKD). The downhill skier lost control of his outside leg during a turn, forcing his inner leg into hip and knee internal rotation with MKD.[35] Throughout the loss of control, hip internal rotation peaked at 50 degrees, with knee internal rotation peaking at 40 degrees, and MKD increasing from 15 degrees to 40 degrees over 100 ms.[35] Knee flexion at time of injury was 65 degrees. All three subjects sustained injuries during a cutting motion, with the majority of weight distributed on a single-leg.[35] Two of the three injuries (basketball and team handball) demonstrated a lack of knee flexion throughout the motion, combined with MKD.[35] In addition, two of the three (team handball and downhill skiing) demonstrated hip internal rotation.[35]

A third study examined videos for body position at time of injury (Krosshaug, *et al.*[36]). They studied the videos of 39 ACL injuries in basketball, both male (n=17) and female (n=22) subjects, with MOI including offense, defense, immediately post-shooting, turnovers, and rebounding.[36] Of the 39 videos, 30 videos showed clear footage of non-contact ACL injuries (13 male and 17 female).[36] Videos were examined at initial contact (IC) with the ground and 50 ms post-IC.[36] Females demonstrated greater average knee flexion at initial contact than males both at IC and 50 ms post (15 degrees at IC/27 degrees at 50 ms vs. 9 degrees at IC/19 degrees at 50 ms).[36] Males and females had similar amounts of MKD at IC, however, females had twice as much MKD 50 ms post compared to males (8 degrees vs. 4 degrees).[36] Nine of the 17 females displayed knee collapse, with two not displaying knee collapse, four not visible and two without consensus.[36] All of these knee collapses were termed *valgus collapse,* being a combination of hip internal rotation, knee valgus and foot external rotation, a definintion similar to MKD.[36] In males, valgus collapse was displayed in only two cases, with 10 cases not displaying valgus collapse.[36] Hip abduction angles were consistent across both genders at both IC and 50 ms, with values ranging from 8-19 degrees.[36]

Overall, these studies show that non-contact ACL injuries occur from a number of different MOI, however they tend to demonstrate a majority of similar number of characteristics: off-balance[34-36], single-leg motions[34-36], hip abduction[34-36], hip internal rotation[34-36], valgus motion at the knee[34-36], knee extended with minimal flexion throughout motion[34-36], internal or external tibial rotation[34-36], and landing heel first or foot-flat[34]. One problem with using footage captured at the time of injury is the lack of knowledge about the athlete prior to injury. In order to acquire information about dynamic movement of athletes prior to injury, research must be conducted prospectively. Unfortunately, it is often difficult to gain funding for large-scale prospective studies due to the potential risk of acquiring no usable data. This means few reliable large-scale prospective studies exist.. However, prospective studies give us valuable insight into the biomechanical risk factors associated with ACL injury by providing us with baseline kinematic (joint angles) and kinetic (joint moments) data at different locations throughout the body, neuromuscular control throughout the body, and medical history of injury, among others. Two of the most important prospective studies in the field of ACL injury research were conducted by Hewett, *et al.*[37] and Zazulak, *et al.*[38]

In 2005, Hewett, *et al.*[37] published data on the first large-scale prospective study involving ACL injury prediction. They examined 205 females who participated in adolescent soccer, basketball or volleyball over the course of four sport seasons.[37] All participants were screened prior to their initial sport seasons collecting kinematic and kinetic data while performing a drop vertical jump.[37] Nine ACL injuries occurred over the course of the observation either during sport practice or during a game. These nine ACL tears were the only injuries that met the requirements of the study[37] Following completion of the study, data from the injured participants was compared with non-injured participants to look for any differences in kinematic or kinetic information.[37]

Injured athletes displayed a significantly greater level of knee abduction and maximum knee displacement than non-injured athletes.[37] Knee abduction of injured athletes was 8.4 degrees more than non-injured athletes while maximum knee displacement of injured athletes was 7.6 degrees more than non-injured participants.[37] In addition, participants who later suffered ACL injury had 10.5 degrees less maximum knee flexion when compared to non-injured participants.[37] Vertical ground reaction force was increased by 20% in athletes who later sustained injury, with significant correlations between knee abduction moments, knee abduction angles, and peak vertical ground reaction forces.[37] Not only were there differences between the injured and non-injured participants, but within the injured participants as well. Injured athletes demonstrated a 6.4 greater side-to-side knee abduction moments between their legs, a difference that was not seen in non-injured participants.[37] Overall, knee abduction moments and angles, both at initial contact and peak values, were significant predictors for risk of ACL injury.[37]

The second study to prospectively examine ACL injury risk was conducted by Zazulak, *et al.*[38] Zazulak, *et al.*[38] prospectively tested 277 athletes, both male and female, and then followed them for three years. The pre-screening consisted of a 45-item questionnaire collecting background demographic information consisting of history of sport participation, history of injury, etc.[38] The study utilized athletes with no history of knee injury.[38] Physical testing consisted of trunk displacement through a weighted release in the flexion, extension and lateral directions with measurements taken at 150 milliseconds and maximum displacement.[38] Participants were restrained at the pelvis to remove any compensatory movements through the hip, knee or ankle.[38] In addition to a restrained pelvis, participants were in a semi-seated position, allowing for participant's most comfortable level of spine adjustment prior to pelvic restraint.[38]

Over the course of three years 25 knee injuries occurred.[38] They were broken down into three categories: knee injuries (KI) (11 female/14 male), 11 knee-ligament injuries (KL) (five female/six male), and six ACL injuries (KACL) (four female/two male).[38] KIs were any injury to the knee, regardless of type, and could include patellar injuries, meniscus tears, etc. KI athletes as well as KACL athletes demonstrated greater trunk displacement than non-injured athletes.[38] Female KL athletes also demonstrated greater maximum trunk displacement when compared to non-injured females.[38] Maximal displacement was removed because 150 milliseconds proved a better predictor of ACL injury.[38]

The final variables included flexion, extension and lateral displacements of the trunk at 150 milliseconds, active proprioceptive repositioning (APR), and history of lower back pain (LBP).[38] Overall, flexion, extension and lateral trunk displacements were the only variables that predicted ACL injury (83% sensitivity and 76% specificity).[38] Between genders, lateral displacement was the only significant predictor of all range of knee injuries (KI, KL, and KACL) in females (100% sensitivity and 72% specificity).[38] APR and history of LBP were also both predictors of knee injury amongst female athletes.[38] In males, statistical significance was only reached in KL injured athletes, with history of LBP being the strongest predictor.[38] One of the limits of the Zazulak study is that the only statistically significant results are specific to the trunk, but still give us a better picture in terms of identifiable risk factors for ACL injury. Both the

Hewett and Zazulak studies helped to illustrate the need to reliably identify movement-based risk factors prior to injury. The ability to identify risk factors during movement will help create more tailored ACL-injury prevention programs and create a more comprehensive rehabilitation protocol in order to prevent either re-injury or contralateral-side injury.

Another method used is manipulation of body position during different motions and seeing how it affected stresses on the ACL. By having participants attempt multiple body positions during a task, researchers are able to use the participant as a control, giving them a greater level of accuracy. Dempsey, *et al.*[39] had 15 males perform a single-leg cutting task using a variety of manipulated body positions (including neutral position, internally or externally rotated foot, wide or narrow foot placement, torso rotation or lateral flexion, etc.) and compared the resulting stresses on the ACL with markers placed at 50 locations around the body.[39] The results showed that compared to a neutral body position, a wide foot stance, the torso leaning opposite the direction of the cut, and the torso rotated away from the direction of the cut all produced much higher valgus or internal rotation stresses on the ACL.[39] Having the foot internally rotated in the direction of the cut, however, decreased stresses on the ACL.[39] Having a wide foot-placement stance during a cutting motion was discovered to increase not only the peak knee valgus stresses, but also peak internal rotation at the knee, when compared to a normal or narrow stance.[40] Combined with decreased knee flexion, a wide stance produces extreme forces around the knee and on the ACL, exposing it to potential for injury.[40] Overall, Dempsey, *et al.*[41]found athletes can display increased MKD or internal rotation at the knee during a cutting motion.

Borotikar, *et al.*[40] found that in addition to all of the other biomechanical risk factors associated with ACL injury risk, fatigue increases risk for females. Following multiple lower extremity exercises, female athletes were required to perform jump-landing tasks which demonstrated a greater amount of hip extension (decrease in hip flexion) as well as an increase in hip internal rotation.[40] Risk factors were also more prominent when performing unplanned cutting tasks versus planned cutting tasks.[40] Unplanned motions are problematic for athletes. Those at risk often do not have time to mentally correct dangerous movements prior to injury. There is not sufficient time for the motor control of the athlete to adjust improper technique to prevent injury. In the case series of 39 athletes, Krosshaug, *et al.*[36] analyzed videos of 17 male and 22 female ACL injuries. They discovered half of the female cases involved secondary contact or pushing of some sort.[36] This implies that a perturbation may not necessarily be the *direct* cause of injury, but may force the athlete into an unplanned movement, compromising their motor control. Many of these non-planned movements are those with the highest risk of injury: jump-landing or cutting tasks.[26]Athletes are thus at more risk when they make unplanned cuts, especially females when combined with fatigue.[26, 36, 40]

Finally, Dempsey, *et al.*[41] found that risks caused by increased MKD or internal rotation at the knee during a cutting motion *could be reduced with an intervention program*. This is key information, because it allows athletes and coaches to be trained in the prevention of injury when an athlete is most predisposed to it. Showing which body positions increase stress on the ACL as well as when stresses are more present emphasizes the use of video screening tools to identify risk factors and the use of injury prevention programs to correct and minimize risk factors.

Recognition of Risk Factors

When considering how best to identify risk factors, 3D motion analysis in a lab setting qualifies as the gold standard. This type of data collection involves a controlled study on a live subject in a laboratory environment. These data include force plates, motion analysis, video analysis, and sometimes supplemental data with fluid samples. For example, blood or saliva allow for tracking hormone levels. Unfortunately, movement analysis using computers is expensive, often costing thousands of dollars to utilize force plates, flock motion analysis, and fluid analysis. In addition, movement analysis requires specialty training for those using the equipment as well as many more specially-trained people and still lacks the practical applications of real-world scenarios. There is a need for a field test that is valid and reliable, easy to use, capable of used on a large scale, and cost effective on large populations.

One type of movement analysis popular in recent years is the video analysis. Video analysis requires very little equipment beyond the cameras and obstacles for the movement, such as a box off which to jump. It mimics 3D motion analysis by using multiple cameras, giving researchers multiple angles to view risk factors, eliminating much of the costly equipment, and is more practical when sampling large populations, as the equipment can be easily transported to different locations with minimal hassle. Video analysis tests do not require the same level of specialty training needed for using 3D motion analysis, lowering the costs of training and increasing availability to a larger population. In addition, video analysis is faster than using 3D motion analysis because the lack of specialty equipment decreases the amount of time needed to test an athlete. Because it is so cost effective, more groups are willing to consider video analysis as an option to screen for potential injury risk. There are multiple different clinical screening tools available for use in risk factor recognition.[42-46]

One example of the practicality of video analysis was when Barber-Westin, *et al.*[43]analyzed the characteristics of jumps and landings in over 1100 athletes aged 9-17. Both genders were recorded doing a single-leg “drop jump”[43], and a single leg hop, to test limb symmetry. The study measured isokinetic strength of both quadriceps and hamstrings. Genders and ages were compared to determine if neuromuscular control of the quadriceps and hamstrings played a role in the increase of ACL injuries following puberty.[43]

In 2009, Padua, *et al.*[42] published the Landing Error Scoring System (LESS), a type of video movement analysis that requires minimal training, minimal expensive equipment, and fewer time or location restraints compared to in-lab biomechanical movement analysis.[42] It was the first video analysis system to accurately capture 3D motion by recording both the front and sagittal planes. The LESS has proven both valid and reliable in recognizing the most common movement risk factors associated with ACL injuries.[42, 46]

The LESS is a major breakthrough for video analysis because it doesn't rely on exact measurements of degrees of flexion and extension. Scorers are not required to use a compass or protractor to measure exact numbers. The LESS relies on simple scores, such as "Peak knee flexion angle: poor < average < excellent". [42] This allows scorers to give a more universal score and reduces discrepancies in both inter- and intra-rater reliability. [42, 46] Scoring the LESS involves determining whether or not someone is demonstrating a risk factor movement. If a participant is displaying a risk factor, they are given a point. The higher someone's score, the more risk factors they demonstrate while performing the LESS, and thus the more at risk they are for an ACL injury. Based on their score, the subject is categorized as having either "poor, moderate, good or excellent" biomechanical control.[42] During initial testing, 36% of women demonstrated "poor" control compared to 23% of men.[42] Only 14% of women demonstrated "excellent" control, compared to 30% of men.[42]

The LESS has proved valuable in assisting correction of movements. Shortly after the publication of the LESS, DiStefano, *et al.*[47] used the LESS as a measurement tool for an ACL prevention and intervention program among young soccer players. The LESS was accurately able to show where intervention was needed to prevent injury. In addition, the LESS was also used as the retest method, to determine whether or not the intervention program was successful at correcting movement errors.[42]

For all of the advantages of the LESS, one of the limitations is that the LESS only analyzes motion when the athlete performs a frontal and vertical motion. Sports do not always occur in two planes; nor do injuries. The simplicity of the LESS is one of the things that make it valuable as a risk assessment tool. However, one of the main components of ACL injury risk is the lateral aspect associated with knee valgus, a planar movement the LESS does not measure.[32, 37, 41] In addition, it does not take into account the single-leg cutting motions often performed in sports, shown to have an increased risk for ACL injury. With such a high number of ACL injuries occurring as non-contact injuries during a single-leg cutting motion, it is important to examine those risky motions.

Statement of Purpose

The high percentage of ACL injuries that occur during a single-leg jump-landing and cutting motion make it vital to find a way of measuring athlete risk. Potential risk factors might not show up as strongly during a different type of movement, such as the frontal and sagittal plane motions analyzed by the LESS. A gap exists in the literature between the frequency of single-leg cutting ACL injuries that occur and the ability to correctly identify the presence of risk factors during the single-leg cutting task. Thus, the goal of this research is to prove that a multi-angle video analysis system can be reliably used to assess ACL injury risk during a single-leg cutting motion.

INTRODUCTION

Injuries are a common risk of participating in sports. Nearly 7 million people are injured playing sports in the United States annually, with 40% of these injuries involving the lower extremity.[48] Of all sports injuries, an estimated 200,000 injuries per year are anterior cruciate ligament (ACL) sprains.[3] At the collegiate level over the last 16 years, the National Collegiate Athletics Association (NCAA) reported a total of 4,800 ACL injuries amongst all sports (an average of 300 per year).[4] Although ACL injuries may not be extremely common, ACL injuries are associated with severe financial and physical consequences.

Injuries to the ACL have high financial costs. ACL-associated repair and rehabilitation costs in the United States average nearly $3 billion annually, with the cost of surgery alone totaling nearly $2 billion annually.[5, 7] The average medical cost associated with surgery and rehabilitation is $15,000 per individual.[3, 7] The physical consequences of ACL injury are even more severe than financial costs and can last for the duration of the individual's lifetime. Following injury, there is an 80-90% chance of developing osteoarthritic changes to the knee within 10-20 years.[17, 49] Using data from the New Zealand national injury registry, Gianotti *et al.*[1] showed that the incidence of ACL surgeries dramatically increase between the ages of 10-14, with the highest rates occurring between the ages of 15-29. The high physical and financial costs put a premium on the ability to reliably identify the risk factors that potentially lead to ACL injury, so they may be corrected and ACL injuries can be prevented.

The majority of ACL injuries (70%) occur due to a non-contact mechanism of injury, such as during single-leg jumping, cutting, or landing motions.[26] Several studies have observed actual mechanisms of ACL injuries through video footage. Their results show that during non-contact single-leg injuries, the injured lower extremity is often extended and abducted away from the body, the knee is in a valgus position, and the foot is externally rotated.[34-36] In addition to observing ACL injuries with video footage, other researchers have used in-vitro studies to determine what stresses the ACL. Markolf, *et al.*[32] manipulated cadavers to load the ACL, ultimately until failure. They discovered that internal or external tibial rotation combined with knee valgus and decreased knee flexion stress the ACL.[32]

Prospective studies support that specific lower extremity movements may predispose an individual for an ACL injury. Hewett, *et al.*[37] examined 205 females performing a jump-landing task. The females were then followed for two years, with nine confirmed ACL injuries. Athletes who suffered ACL injury had greatly altered body posture including poor trunk neuromuscular control, greater valgus motion at the knee, and increased hip abduction compared to non-injured athletes.[37] Zazulak, *et al.*[38] also performed a prospective study looking at risk factors for ACL injury, however, this study focused exclusively on risk factors associated with the trunk. Their results indicated that the lack of neuromuscular control with flexion, extension, and lateral trunk flexion were all predictors of ACL injury in females, but not males.[38] Therefore, ACL injury prevention may be possible by screening individuals for high-risk movements and teaching individuals to avoid these body positions.

ACL injury prevention programs may be more effective if we can successfully identify movement-based risk factors.[50, 51] In addition, risk factor identification allows us to ensure modification of risk factors post-ACL reconstruction, which may help to prevent repeat or contralateral side injuries. Three-dimensional motion analysis is considered the gold standard of movement analysis, however, these systems are expensive, require specialty training, and are not practical for use in the field.[37, 39] Another method to evaluate movement-based risk factors for injury is a clinical screening tool that utilizes two-dimensional video analysis. Video analysis is cost effective, requires minimal training, and is able to be used on a large scale in the field.[42, 46] Clinical screening tools using video analysis are numerous and varied in methods.[42-45, 52] One video analysis system that has proven both valid against three-dimensional motion analysis and reliable is the Landing Error Scoring System (LESS).[42, 46]

The LESS is a simple field test involving a double-leg jump landing that can be used on a large scale to correctly identify motions are considered to be high risk for ACL injury.[42, 46] However, the LESS is not capable of identifying movement during single-leg activities, such as a cutting task. This is true with most movement-based video analysis systems, as they utilize a double-leg landing.[42-45] Currently, there is a gap in the literature in the realm of ACL injury risk factor video analysis during a direction-change or cutting motion with a single leg. Consequently, we created the Cutting Screening Tool (CST) to fill that gap in the literature of identifying ACL risk factors during single-leg cutting movements. The purpose of this study is to determine if the CST is a reliable clinical screening tool to evaluate movement technique during a single-leg cutting motion.

METHODS

Experimental Approach to the Problem

We used a repeated measures design to evaluate the intra-rater reliability of the CST and a cross-sectional design to evaluate the inter-rater reliability of the CST. A random subsample of video data from a previous study that evaluated cutting biomechanics of high school athletes were graded using the CST. Participants graded 20 individuals performing 3 trials of the cutting task one time, and graded a random subsample of 5 individuals a second time at least 3 days after completing the first round of grading.

Subjects

Participants were recruited via email using the clinical instructor and the alumni list-serves from the University of Connecticut Athletic Training Education Program. Athletic trainers were eligible to participate if they were currently licensed and practicing clinically in the state of Connecticut. Participants were excluded from the study if they had received formal training in clinical movement screening tools, such as the Landing Error Scoring System (LESS). Participants were recruited until three participants with greater than 10 years of clinical experience post-certification and three participants with less than 5 years of experience had volunteered to participate in this study.

Rating candidates' years of experience was selected on the theory that raters with less than five years' experience would rely more heavily on their formal education whereas participants with more than 10 years' experience would be more likely to rely on field experience and the knowledge gained as a result.

Raters with less than five years of clinical experience were chosen because their formal education experience would be more current. Formal education with regards to ACL injuries has drastically changed over the last ten years (citation?). As a result, participants with less than five years of clinical experience would have a more up to date knowledge of current best practices. However, they would not necessarily have the exposure to a large population of athletes in their shorter time in a clinical setting, making it potentially difficult for them to intuitively identify

ACL injury risk factors. For these reasons, raters with less than five years of clinical experience were classified as “Novice”.

While not having the most up to date formal education on ACL injuries, raters with more than 10 years’ experience would have a stronger practical base at identifying ACL risk factors. More years of clinical experience potentially made them faster at intuitively assessing and analyzing the movement patterns of athletes, allowing them to understand and incorporate the training faster than someone with less clinical experience. For these reasons, raters with more than 10 years of clinical experience were classified as “Veteran”.

Procedures

High school athletes from a previous study successfully completed three trials of a side-step cutting task. During the task, these participants were videotaped using two standard digital video cameras (Sony Products, Park Ridge, NJ) which were positioned directly in front of the participant and to the right side of the participant to capture both frontal and sagittal plane images. The sidestep cutting task required participants to jump forward from a 30-cm high box using their non-dominant limb a distance of half their body height, land in a target area on their dominant limb, and perform a 60 degree sidestep cut toward their non-dominant limb (Figure 2). The dominant limb was defined as the limb used to kick a ball for greatest distance.

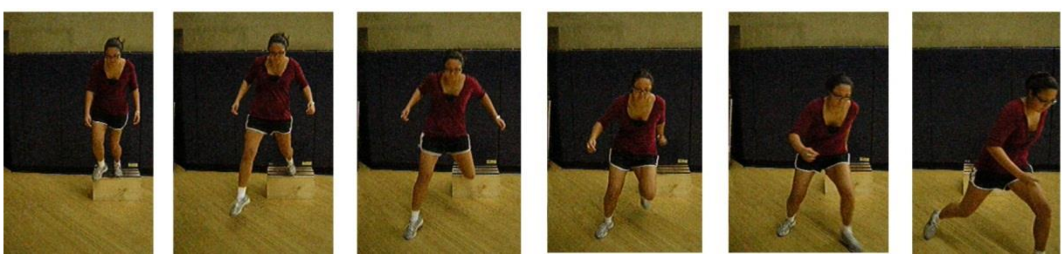

Figure 2. The standardized sidestep cutting task.

The Cutting Screening Tool (CST)

The Cutting Screening Tool is designed to be a clinical tool to evaluate movement during a side-step cutting task. The CST uses dichotomous scoring (“Error” or “No Error”) to evaluate specific movements. A total score is computed based on the number of “errors” an individual makes while performing the cutting task. An overall low score indicates better movement technique during the cutting task. Raters are instructed to default to “No Error” if in question about whether or not an individual possesses a specific error.

Training Session

The participants attended a single training session for the CST, which lasted between one and two hours. During the training session, participants were provided with a short background on ACL injuries and the rationale for evaluating lower extremity movement during a cutting task. In addition, participants were taught to watch for the specific lower extremity movements on the CST, such as position of the foot and tri-planar motion of the trunk, hips, and knee. Participants received copies of the operational definitions (Table 1) for the different errors and the CST grading sheet (Appendix A). Participants viewed examples of each error in still-pictures, and asked questions throughout the presentation to clarify how the errors were identified and classified. Throughout the presentation, participants were allowed to take any notes they felt necessary to help them correctly identify the different errors during grading.

ITEM	OPERATIONAL DEFINITION	CAMERA VIEW
Foot placement	Once the entire foot is in contact with the ground, visualize a line down from the greater trochanter. If the line falls inside the foot, placing the foot outside greater trochanter (wide), score ERROR. If the test foot is not outside the greater trochanter, score NO ERROR. If the test foot is internally or externally rotated, grade based on heel placement.	Front
Foot rotation	If foot is externally rotated more than 30 degrees, score ERROR. If the foot is not externally rotated more than 30 degrees at initial contact, score NO ERROR.	Front
Hip drop	If hip drops from initial contact to maximal knee flexion, score ERROR. If it does not, score NO ERROR.	Front
Lateral trunk Displacement	If the trunk flexes laterally away from the direction of the cut, score ERROR. If the trunk stays vertical or flexes in the direction of the cut, score NO ERROR.	Front
Medial knee Displacement	At the point of maximal medial knee position, visualize a line straight down from the center of the patella. If the line runs through the great toe or is medial to the great toe, score ERROR. If the line is lateral to the great toe, score NO ERROR.	Front
Plantar flexion at initial contact	If the foot lands heel to toe or with a flat foot, score ERROR. If the foot lands toe to heel, score NO ERROR.	Side
Knee flexion Displacement	If the knee does not flex more than 45 degrees during the cut, score ERROR. If the knee flexes more than 45 degrees, score NO ERROR.	Side
SMALL trunk flexion displacement	If the trunk does not flex from initial contact to maximum knee flexion, score ERROR. If the trunk does flex, score NO ERROR.	Side
EXCESSIVE trunk flexion displacement	If the trunk flexes past parallel with the lower leg, score ERROR. If the trunk appears parallel or less with the lower leg, score NO ERROR.	Side
Joint displacement	Watch the sagittal plane motion at the hip and knee from initial contact to max knee flexion angle. If the subject subject goes through large displacement of the trunk, hip, and knee, score SOFT. If the subject goes through some trunk, hip, and knee displacement but not a large amount, then AVERAGE. If the subject goes through very little, if any trunk, hip, and knee displacement, then STIFF.	Side
Overall impression	Score EXCELLENT if the subject displays a soft landing and no frontal or transverse plane motion at the knee. Score POOR if the subject displays a stiff landing and some frontal or transverse plane motion OR large frontal and transverse plane motion at the knee. All other landings, score AVERAGE.	Side, Front

Table 1. Operational Definitions for the Cutting Screening Tool.

Following the presentation, participants practiced grading sample videos with the primary investigator using the CST. Participants were shown a sample video and attempted to grade the first trial with minimal feedback by the primary investigator. Following the participant's first grading trial the primary investigator gave any further clarifications as needed to assure that participants understood the grading procedures. A second trial of the same subject was then shown. Additional necessary feedback was given by the primary investigator after the participant had finished. After two trials, a third trial of a different subject was shown, allowing participants to practice on multiple subjects. At this time, the participants asked any remaining questions they had for clarification. After the training session, they were given a packet containing all the necessary supplies for completing the grading. Necessary supplies included the operational definitions, grading sheets, a disc containing all the videos to grade, a comments and feedback sheet about using the CST, and a baseline questionnaire for demographic information. Participants were instructed to begin the initial grading process within one week of completing the initial training session, meaning that all grading would be completed within two weeks of training.

Each participant graded the same 20 high school participant videos. Participants were instructed to review and grade each trial independently of other trials. This was to prevent participants from becoming biased towards or against certain errors throughout the grading process and ensure the reliability of the scores for each trial. Participants were instructed to grade the errors in order of appearance on the grading sheet. Even if an error was noticed earlier, they were instructed to ignore it until they came to that particular item on the scoring sheet. This created a consistent flow amongst the trials. If participants felt the order of grading was difficult or confusing, they were instructed to indicate such on the comments and feedback sheet.

If participants became confused as to what qualifies as an error during grading, they were told to refer back to the operational definitions. The participants graded the videos on their own time. Participants were told that once grading began, if they found within one or two videos that they were having difficulty with a specific part of the CST, to contact the primary investigator for clarification. Once the 20 individuals (60 videos) were graded a first time, the participants were instructed to wait a minimum of three days, but no more than one week, before grading five individuals' videos a second time. The five sets of videos were randomly chosen from the initial 20 sets of videos. These videos were used to evaluate intra-rater reliability.

Data Reduction and Analyses

Following collection of completed forms, data was entered into a spreadsheet (Microsoft Excel, Microsoft Corp., Redmond, Washington). To compare inter-rater reliability multiple different methods, including intra-class correlation coefficient (ICC), Cohen's kappa test statistic, Fleiss' kappa test statistic, and Bland-Altman plots were used. To compare intra-rater reliability the ICC and Bland-Altman plots were used.

Inter-rater Reliability

Data were input and calculated using SPSS (Version of SPSS - 19) (SPSS Inc, Chicago, IL, USA) and Excel. ICC, Cohen's kappa, and Fleiss' kappa results are typically assigned descriptions of the values. This study chose to interpret the values as Excellent (1.00-0.801), Good (0.80-0.601), Moderate (0.60-0.401), Poor (0.40 >). ICC scores were calculated for total scores of the CST. In addition to group total scores, we also calculated an ICC for total scores when separated by experience (Novice and Veteran).

Inter-rater reliability of individual items were evaluated using the Cohen's kappa test statistic and Fleiss' kappa statistic. Cohen's kappa statistics use a dichotomous formula that includes random chance agreement, making it a stronger statistic than percentage agreement alone. Fleiss' kappa analyzes data amongst multiple raters, making it a valuable tool when looking at the overall comparisons between raters. In addition, Fleiss' kappa doesn't require multiple data point results, giving only a single result.

When calculating the scores of the final two items of the CST, we were unable to use a standard kappa statistic initially because of the nature of the scoring. In the final two items of the CST, scores are ranked 0, 1, or 2 instead of 0 or 1 as in the rest of the test. As a result, we separated the scores into two categories. For #10, Joint Displacement, scores were separated into "STIFF" and "SOFT". When calculating "STIFF", any score below 2 was replaced with a 0, while a score of 2 was replaced with a 1. To calculate "SOFT", any score greater than 0 was replaced with a 1, while any score of 0 remained the same. To calculate #11 (Overall Impression), the same method was used for separating the scores into "POOR" and "EXCELLENT". When calculating "POOR", a score below 2 was changed to a 0, while scores of 2 were changed to 1. For "EXCELLENT", any score above 0 was changed to 1.
One limitation of kappa statistics occurs when there is a high amount of chance agreement. Chance agreement is the idea that if the raters simply guessed on their scoring, there would still be a percentage agreement in the results. Kappa statistics account for this, but if the risk of chance agreement is high, it can create skewed results. If this occurs, the resulting kappa statistic will be low, and percentage agreement is more effective. This is found in previous testing of binary clinical screening tools, and applies to both Cohen's kappa and Fleiss' kappa.[46]

As a result, we included both kappa statistic scores as well as the number of agreements between raters. Due to the high number of chance agreement in the Cohen's kappa statistic, certain numbers were replaced with a *, indicating a 1.00 chance agreement. When calculating the average score for the Cohen's kappa statistic, scores with a * were replaced with a 1.00 as the correlation is extremely high for these numbers.

In addition to Kappa statistics, we also tested inter-rater reliability of individual items using a Bland-Altman plot. A Bland-Altman plot is a pairwise comparison of data. Creation of a Bland-Altman uses the difference from the mean of two data points. Data points are the total number of errors scored by each rater for each individual item. Bland-Altman gives us visual representation of variability amongst the overall percent scored positive or negative for the different raters. Bland-Altman interpretations were relative to the graphs of the other individual items. Using Bland-Altman plots, we are able to single out specific tests with increased variability or specific raters with increased variability. When reading a Bland-Altman plot, the X-axis represents the mean of two raters when scoring an item and the Y-axis represents the difference of the two raters from the mean.

For example, if one rater scores an error 10/20 (0.5), and the second rater scores 12/20 (0.60), the mean would be 11/20 (0.55 on the X-axis). Then, the score of the second rater (0.60) would be subtracted from the scores of the first rater (0.50), placing the point on the Y-axis at (-0.10). The final data point on the graph would be (0.55,-0.10). When we plot multiple data points on the graph, we see detailed results. If a single rater's results are constantly different from the results of other raters measureable through distance from 0.00 on the Y-axis. If the data points remain close to 0.00 on the Y-axis, it shows difference between raters' results are minimal. When data points spread out on both the X and Y-axes, it shows a general lack of reliability of a specific item. For the individual item Bland-Altman plot, Joint Displacement and Overall Impression were both broken down into two separate categories using the same methods as the Kappa statistics.

Intra-rater Reliability

When considering intra-rater reliability, we used both an ICC and Bland-Altman. ICC was calculated for overall intra-rater reliability as well as overall intra-rater reliability when divided by level of experience. We removed one rater from a majority of intra-rater calculations and graphs due to a discrepancy in the subject videos graded for the repeat trials. A different set of videos were completed for the repeat trials. To avoid contamination of scores, V3 was removed. Bland-Altman plots were created based on total scores of raters for each of the five subject videos that were graded on repeat trials. Each graph was then color-coded to show intra-rater reliability of each video based on experience level. This was done to support intra-rater reliability results from the ICC.

RESULTS

All six participants completed the training session and grading requirements. Completion time for the graded sheets varied from one to four weeks. Three individuals had been certified less than five years (3 Females; Avg. Years Certified = 3.17±1.65; Avg. Years Licensed in Connecticut = 2.67±1.25; Number attended Accredited Program = 3; Avg. Years in Settings: College = .67±.94, High School = 2.33±1.55, Clinic = 1.5±2.12) were classified as “Novice”(N1-N3). Three individuals had been certified more than 10 years (2 Males, 1 Female; Avg. Years Certified = 26.17±6.64; Avg. Years Licensed in Connecticut = 11.17±11.17; Number attended Accredited Program = 0, Number attended Internship = 3; Avg. Years in Settings: College = 21.67±11.47, High School = 6.33±8.96, Clinic = 5±6.38, Professional = 0.67±0.94) were classified as “Veteran” (V1-V3).

Inter-Rater Reliability

Inter-rater Reliability of Average Total CST Score: Intra-class Correlation Coefficient

The overall inter-rater reliability between all 6 raters was poor ($ICC_{(2,1)} = 0.32$, SEM = 1.42). Table 2 presents the pairwise inter-rater reliability values, which ranged from poor to good (range: $ICC_{(2,1)}$= 0.02-0.64). The highest inter-rater correlation was seen between V2 and R2 or V2 and V1. The lowest inter-rater correlation was seen between R3 and V3. Inter-rater reliability when comparing the Novice and Veteran group was good ($ICC_{(2,1)} = 0.64$, SEM = 0.69).

INTER-RATER RELIABILITY					
	N 2	**V 1**	**N 3**	**V 2**	**V3**
N 1	ICC=0.27 SEM=1.03	ICC=0.40 SEM=1.24	ICC=0.12 SEM=1.15	ICC=0.46 SEM=1.25	ICC=0.06 SEM=1.18
N 2		ICC=0.38 SEM=1.26	ICC=0.37 SEM=0.95	ICC=0.64 SEM=1.03	ICC=0.39 SEM=0.94
V 1			ICC=0.53 SEM=1.10	ICC=0.64 SEM=1.04	ICC=0.12 SEM=1.50
N 3				ICC=0.36 SEM=1.38	ICC=0.02 SEM=1.17
V 2					ICC=0.19 SEM=1.55

Table 2. Inter-rater reliability based on total scores.

Inter-rater Reliability of Individual Items: Cohen's Kappa Statistic

Results from inter-rater reliability of individual items are found in Tables 3-15. Scores with 1.00 chance agreement are marked with a * due to the inability to divide by 0. Negative scores indicate a higher random chance agreement than actual agreement.

Foot Placement: Average Cohen's kappa statistic score for Foot Placement was good at 0.64, and Fleiss' kappa statistic score was 0.51. The range of agreement amongst raters was 15-19 out of 20. The highest correlation was seen between V2 vs. V3, with a score of excellent at 0.90. The lowest correlation was seen between V1 vs. V3 or R3 vs. V3 with a moderate score of 0.50.

Foot Rotation: The average Cohen's kappa statistic score was good at 0.74, and a Fleiss' kappa statistic score of -0.11. The range of agreement was 17-20 out of 20. Foot rotation had the highest correlation between V1 vs. V3 and was excellent (*) 1.00. The lowest score was seen between N1 vs. V1 or N1 vs. V3 and was poor at 0.27.

Hip Drop: The average score for hip drop was 0.17. The range of agreement was 5-18 out of 20. The highest score was seen between N1 vs. N2, and was good at 0.60. Cohen's kappa score was lowest between N3 vs. V3, and was poor at -0.13, with a Fleiss' kappa statistic score of 0.01, which was also poor.

Lateral Trunk Displacement: The average Cohen's kappa statistic for lateral trunk displacement was 0.47 with a range of agreement 10-19 out of 20. Fleiss' kappa statistic score was 0.31. The highest correlation was seen between N3 vs. V2, and was excellent at 0.90. The lowest correlation was seen in N1 vs. N2, and was poor at -0.05.

Medial Knee Displacement: The average Cohen's kappa statistic was 0.32, and a Fleiss' kappa statistic score of 0.07. The range of agreement was betweeN17-20 out of 20. The highest correlation was seen in N2 vs. N3 and was excellent at 1.00 (*). The lowest correlation was seen in N2 vs. V3 or N3 vs. V3 at -0.20.

Plantar Flexion at Initial Contact: The average Cohen's kappa was 1.00, and a Fleiss' kappa statistic score of -0.10. Every inter-rater comparison had 20 agreements. Highest correlation seen in N1 vs. V2, N1 vs. V3, and V2 vs. V3 at 1.00 (*) and is considered excellent. The lowest correlation was 1.00, was present in all other correlations, and qualifies as excellent.

Knee Flexion Displacement: The average correlation of Cohen's kappa was 0.36 with a range of agreement betweeN15-20 out of 20. Fleiss' kappa statistic score was 0.04. The highest correlation in Knee Flexion Displacement was in N1 vs. N2 with 1.00, and is excellent. Lowest correlations were in N1 vs. V1, N1 vs. V3, N2 vs. V3, V1 vs. N3, and N3 vs. V3. All had correlations of 0.10, considered poor.

Small Trunk Flexion Displacement: The average correlation of Cohen's kappa was 0.19 with a range of agreement between 5-18 out of 20. Fleiss' kappa statistic had a score of 0.01. Highest correlations in trunk flexion displacement seen in N2 vs. V2 with 0.72 and qualifies as good. The lowest correlation was found in V2 vs. V3 with a score of -0.09 and is considered poor.

Excessive Trunk Flexion Displacement: The average agreement for the Cohen's kappa was 0.54, with a Fleiss' kappa statistic score of 0.33. The range of agreement was 19-20 out of 20. Highest correlation seen in N1 vs. N2, N1 vs. N3, N2 vs. V2, N2 vs. V3, V1 vs. N3, N3 vs. V2, and N3 vs. V3. These correlations all had a score of 1.00 and are considered excellent. The lowest correlation seen in V1 vs. V3 is -0.50 and is considered poor.

Joint Displacement Stiff: The average agreement of Cohen's kappa was 0.10 with range of agreement betweeN14-20 out of 20. Fleiss' kappa statistic score was 0.65. The highest agreement was in N2 vs. V3 at 1.00 (*) and is considered excellent. The lowest correlation was N3 vs. V3 at -1.00 and is considered poor.

Joint Displacement Soft: The average Cohen's kappa was 0.20 with a range of agreement 12-20 out of 20. Fleiss' kappa statistic had a score of 0.65. Highest correlations seen in N1 vs. N2, N1 vs. N3, and N2 vs. N3 at 1.00 (*) and is considered excellent. Lowest correlation seen in V2 vs. V3 at -0.25 and is considered poor.

Overall Impression Poor: The average correlation was 0.10 with a Fleiss' kappa score of 0.60. The range of agreement was 8-20 out of 20. Highest correlation seen in N2 vs. V3 at 1.00 (*), considered excellent. Lowest correlation was N1 vs. V1 at -0.17 and is poor.

Overall Impression Excellent: The average Cohen's kappa statistic score was 0.43, and a Fleiss' kappa statistic score of 0.60. The range of agreement was betweeN17-20 out of 20. Highest correlation in N1 vs. N2, N1 vs. N3, N1 vs. V2, N2 vs. N3, N2 vs. V2, and N3 vs. V3 at 1.00 (*), qualifying as excellent. The lowest correlation in V1 vs. V3 at -0.17, qualifying as poor.

Kappa Statistic	Agreement
<0	Poor
0-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Excellent

Table 3. Interpretation of kappa statistic results

FOOT PLACEMENT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		0.71 NA: 17/20	0.62 NA: 16/20	0.71 NA: 17/20	0.60 NA: 16/20	0.70 NA: 17/20
N2			0.66 NA: 17/20	0.69 NA: 17/20	0.70 NA: 17/20	0.60 NA: 16/20
V1				0.67 NA: 17/20	0.52 NA: 15/20	0.50 NA: 15/20
N3					0.61 NA: 16/20	0.50 NA: 15/20
V2						0.90 NA: 19/20

Table 4. Inter-rater reliability of foot placement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

FOOT ROTATION						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		0.30 NA: 17/20	0.27 NA: 17/20	0.30 NA: 17/20	0.30 NA: 17/20	0.27 NA: 17/20
N2			1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20
V1				1.00 NA: 20/20	1.00 NA: 20/20	* NA: 20/20
N3					1.00 NA: 20/20	1.00 NA: 20/20
V2						1.00 NA: 20/20

Table 5. Inter-rater reliability of foot rotation. NA: Number of Agreements (either error present or not present) out of 20 subjects.

HIP DROP

	N1	N2	V1	N3	V2	V3
N1		0.60 NA: 16/20	0.15 NA: 12/20	-0.07 NA: 9/20	0.28 NA: 13/20	0.16 NA: 12/20
N2			-0.03 NA: 10/20	-0.08 NA: 9/20	0.49 NA: 15/20	0.08 NA: 11/20
V1				-0.05 NA: 5/20	0.26 NA: 15/20	0.48 NA: 18/20
N3					0.15 NA: 10/20	-0.13 NA: 5/20
V2						0.28 NA: 15/20

Table 6. Inter-rater reliability of hip drop. NA: Number of Agreements (either error present or not present) out of 20 subjects.

LATERAL TRUNK DISPLACEMENT

	N1	N2	V1	N3	V2	V3
N1		-0.05 NA: 12/20	0.38 NA: 16/20	0.06 NA: 10/20	0.13 NA: 11/20	0.57 NA: 18/20
N2			0.33 NA: 14/20	0.61 NA: 16/20	0.50 NA: 15/20	0.40 NA: 15/20
V1				0.42 NA: 14/20	0.51 NA: 15/20	0.47 NA: 16/20
N3					0.90 NA: 19/20	0.23 NA: 12/20
V2						0.31 NA: 13/20

Table 7. Inter-rater reliability of lateral trunk displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

MEDIAL KNEE DISPLACEMENT

	N1	N2	V1	N3	V2	V3
N1		0.40 NA: 19/20	0.28 NA: 17/20	0.40 NA: 19/20	0.28 NA: 17/20	0.02 NA: 17/20
N2			0.33 NA: 18/20	* NA: 20/20	0.33 NA: 18/20	-0.20 NA: 18/20
V1				0.33 NA: 18/20	0.41 NA: 17/20	0.52 NA: 18/20
N3					0.33 NA: 18/20	-0.20 NA: 18/20
V2						0.52 NA: 18/20

Table 8. Inter-rater reliability of medial knee displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

PLANTAR FLEXION AT INITIAL CONTACT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20	* NA: 20/20	* NA: 20/20
N2			1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20
V1				1.00 NA: 20/20	1.00 NA: 20/20	1.00 NA: 20/20
N3					1.00 NA: 20/20	1.00 NA: 20/20
V2						* NA: 20/20

Table 9. Inter-rater reliability of plantar flexion. NA: Number of Agreements (either error present or not present) out of 20 subjects.

KNEE FLEXION DISPLACEMENT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		1.00 NA: 20/20	0.10 NA: 17/20	0.40 NA: 18/20	0.50 NA: 19/20	0.10 NA: 17/20
N2			0.25 NA: 17/20	0.50 NA: 18/20	0.50 NA: 19/20	0.10 NA: 17/20
V1				0.10 NA: 15/20	0.14 NA: 16/20	0.64 NA: 18/20
N3					0.36 NA: 17/20	0.10 NA: 15/20
V2						0.57 NA: 18/20

Table 10. Inter-rater reliability of knee flexion displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

SMALL TRUNK FLEXION DISPLACEMENT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		-0.01 NA: 11/20	0.26 NA: 12/20	0.00 NA: 11/20	0.25 NA: 13/20	-0.11 NA: 10/20
N2			-0.01 NA: 5/20	0.48 NA: 18/20	0.73 NA: 16/20	0.22 NA: 17/20
V1				-0.03 NA: 5/20	0.14 NA: 9/20	0.32 NA: 10/20
N3					0.38 NA: 16/20	0.27 NA: 17/20
V2						-0.09 NA: 13/20

Table 11. Inter-rater reliability of small trunk flexion displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

EXCESSIVE TRUNK FLEXION PLACEMENT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		0.38 NA: 19/20	1.00 NA: 20/20	1.00 NA: 20/20	0.38 NA: 19/20	-0.02 NA: 19/20
N2			0.25 NA: 19/20	0.48 NA: 19/20	1.00 NA: 20/20	1.00 NA: 20/20
V1				1.00 NA: 20/20	0.25 NA: 19/20	-0.50 NA: 19/20
N3					1.00 NA: 20/20	1.00 NA: 20/20
V2						-0.09 NA: 20/20

Table 12. Inter-rater reliability of excessive trunk flexion displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

JOINT DISPLACEMENT STIFF						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		0.00 NA: 15/20	0.20 NA: 14/20	0.73 NA: 16/20	0.50 NA: 17/20	0.00 NA: 15/20
N2			0.00 NA: 15/20	0.00 NA: 19/20	0.00 NA: 18/20	* NA: 20/20
V1				-0.09 NA: 14/20	0.17 NA: 15/20	0.00 NA: 15/20
N3					-0.07 NA: 17/20	-1.00 NA: 18/20
V2						0.00 NA: 18/20

Table 13. Inter-rater reliability of poor joint displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

JOINT DISPLACEMENT SOFT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		* NA: 20/20	0.00 NA: 19/20	* NA: 20/20	0.00 NA: 16/20	0.00 NA: 16/20
N2			0.00 NA: 19/20	* NA: 20/20	0.00 NA: 16/20	0.00 NA: 16/20
V1				0.00 NA: 19/20	0.35 NA: 17/20	-0.09 NA: 15/20
N3					0.00 NA: 16/20	0.00 NA: 16/20
V2						-0.25 NA: 12/20

Table 14. Inter-rater reliability of excellent joint displacement. NA: Number of Agreements (either error present or not present) out of 20 subjects.

OVERALL IMPRESSION POOR						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		0.00 NA: 15/20	-0.17 NA: 13/20	0.00 NA: 11/20	0.00 NA: 9/20	0.00 NA: 15/20
N2			0.00 NA: 18/20	0.00 NA: 12/20	0.00 NA: 8/20	* NA: 20/20
V1				0.29 NA: 14/20	0.14 NA: 10/20	0.00 NA: 18/20
N3					0.23 NA: 12/20	0.00 NA: 12/20
V2						0.00 NA: 8/20

Table 15. Inter-rater reliability of poor overall impression. NA: Number of Agreements (either error present or not present) out of 20 subjects.

OVERALL IMPRESSION EXCELLENT						
	N1	**N2**	**V1**	**N3**	**V2**	**V3**
N1		* NA: 20/20	0.00 NA: 19/20	* NA: 20/20	* NA: 20/20	0.00 NA: 18/20
N2			0.00 NA: 19/20	* NA: 20/20	* NA: 20/20	0.50 NA: 19/20
V1				0.00 NA: 19/20	0.00 NA: 19/20	-0.07 NA: 17/20
N3					* NA: 20/20	0.00 NA: 18/20
V2						0.00 NA: 18/20

Table 16. Inter-rater reliability of excellent overall impression. NA: Number of Agreements (either error present or not present) out of 20 subjects.

Inter-rater Reliability of Individual Items: Bland-Altman Test

A Bland-Altman graphical representation of inter-rater comparisons for every item on the CST is seen in Figure 3. Note: due to a clerical error with data entry, Veteran raters V1-V3 are scored in Bland-Altman graphs as E1-E3, but the data points are the same.

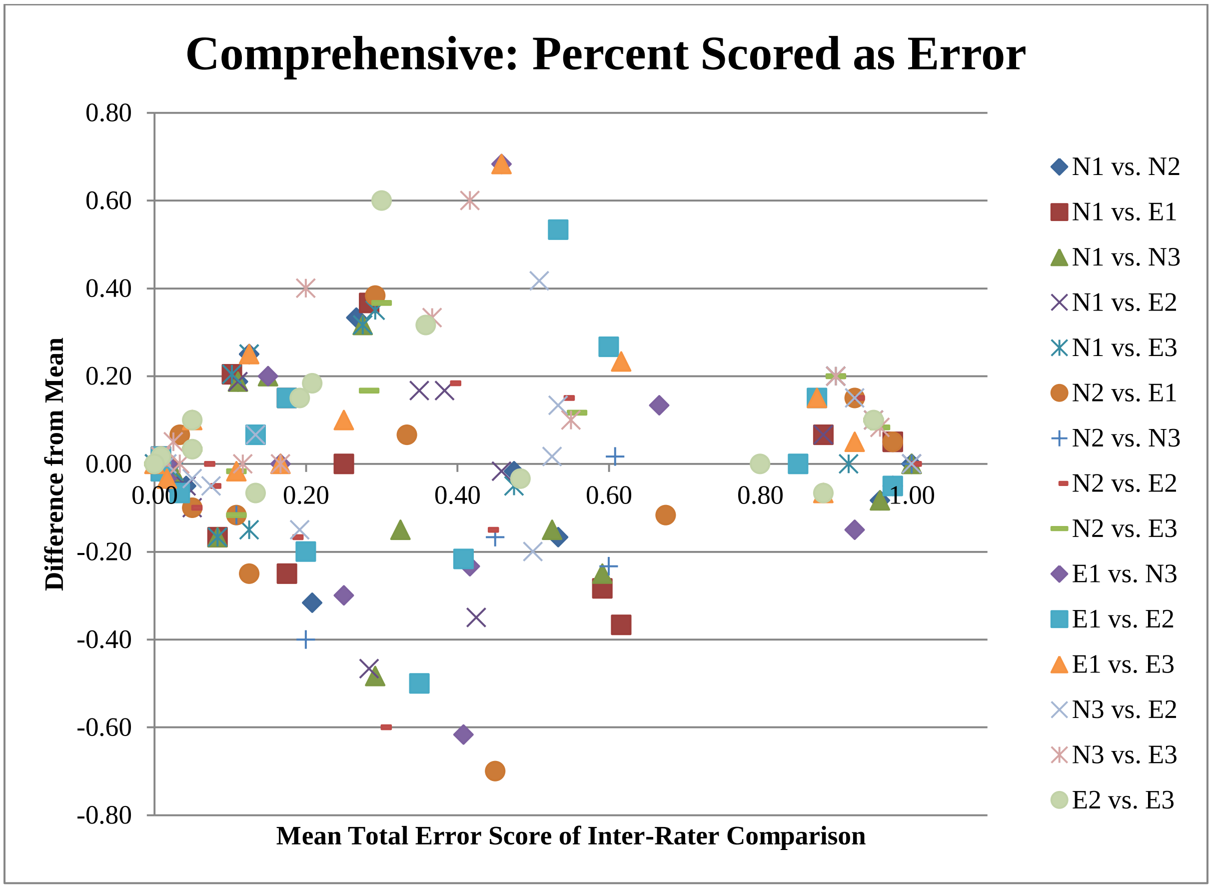

Figure 3. Comprehensive display of inter-rater comparisons for every item on the CST.

Figure 3 shows the results from the inter-rater reliability of every item on the CST. For this graph, the distance of a data point from 0.00 on the Y-axis is important. An excessive distance from 0.00 on the Y-axis shows increased variability in the scores provided by the raters. A highly-correlated Bland-Altman graph would show a majority of data points around 0.00 on Y-axis. In addition, data points occasionally overlap, appearing as though certain data points are not present. Bland-Altman graphs for each individual item are found in Figures 4-16.

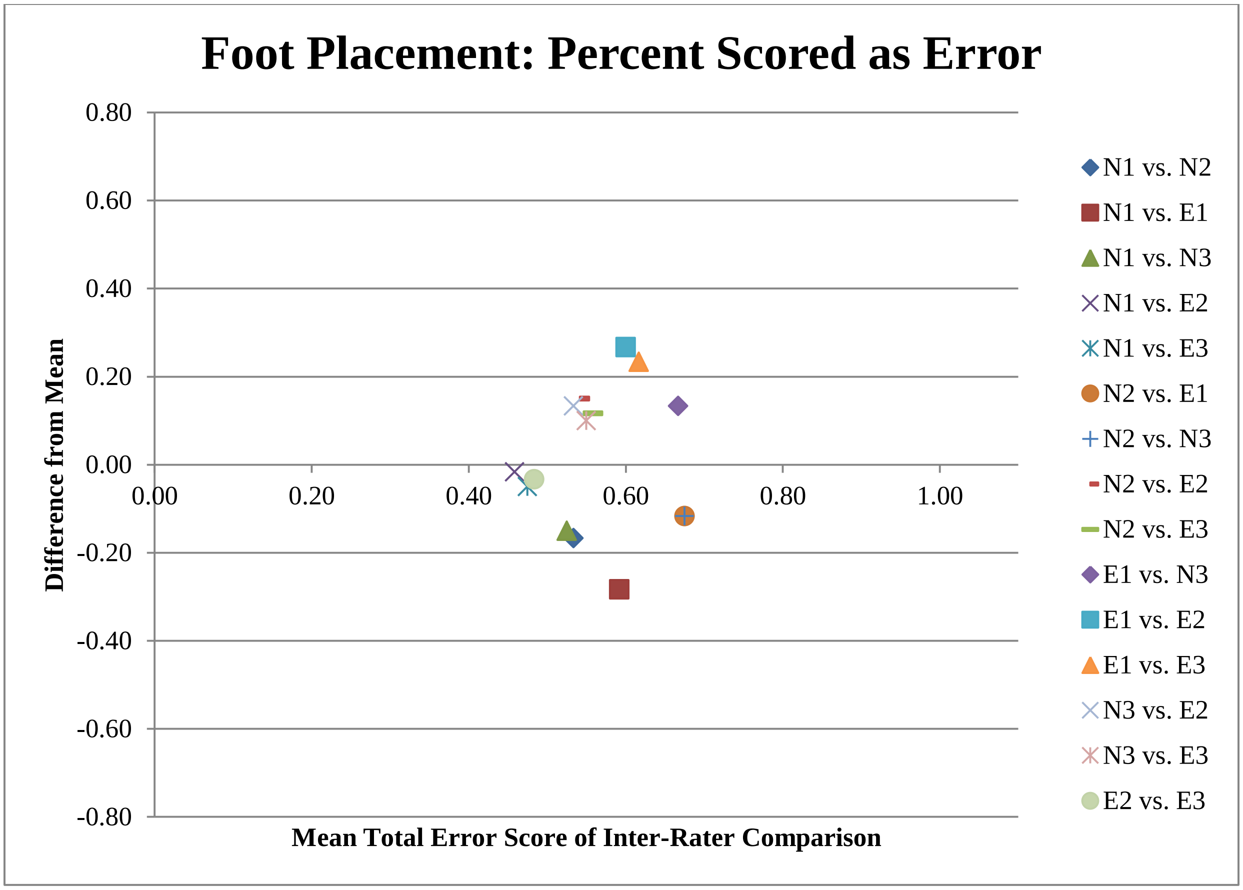

Figure 4. Bland-Altman graph of Foot Placement. Note: Overlap of data points may cause appearance of missing data points.

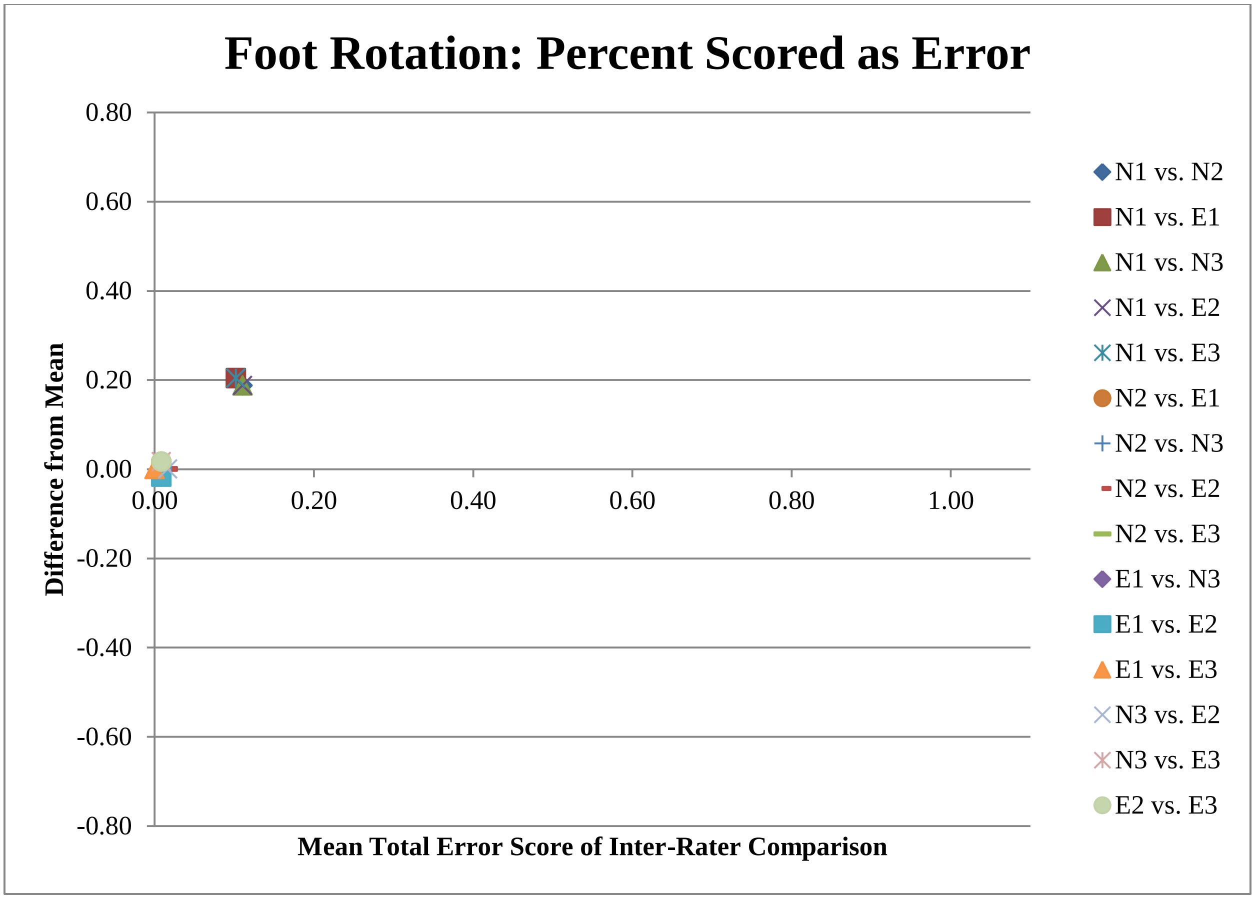

Figure 5. Bland-Altman graph of Foot Rotation. Note: Overlap of data points may cause appearance of missing data points.

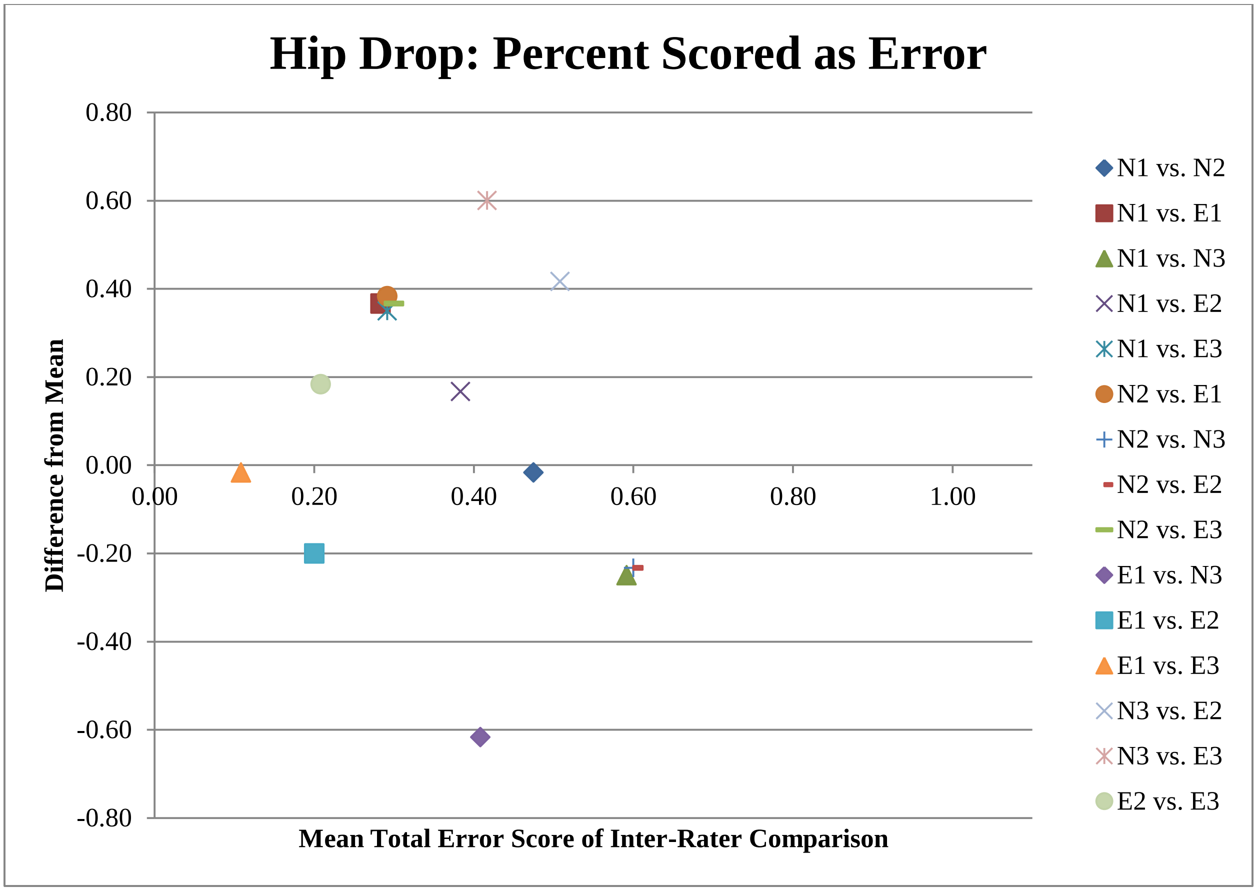

Figure 6. Bland-Altman of Hip Drop. Note: Overlap of data points may cause appearance of missing data points.

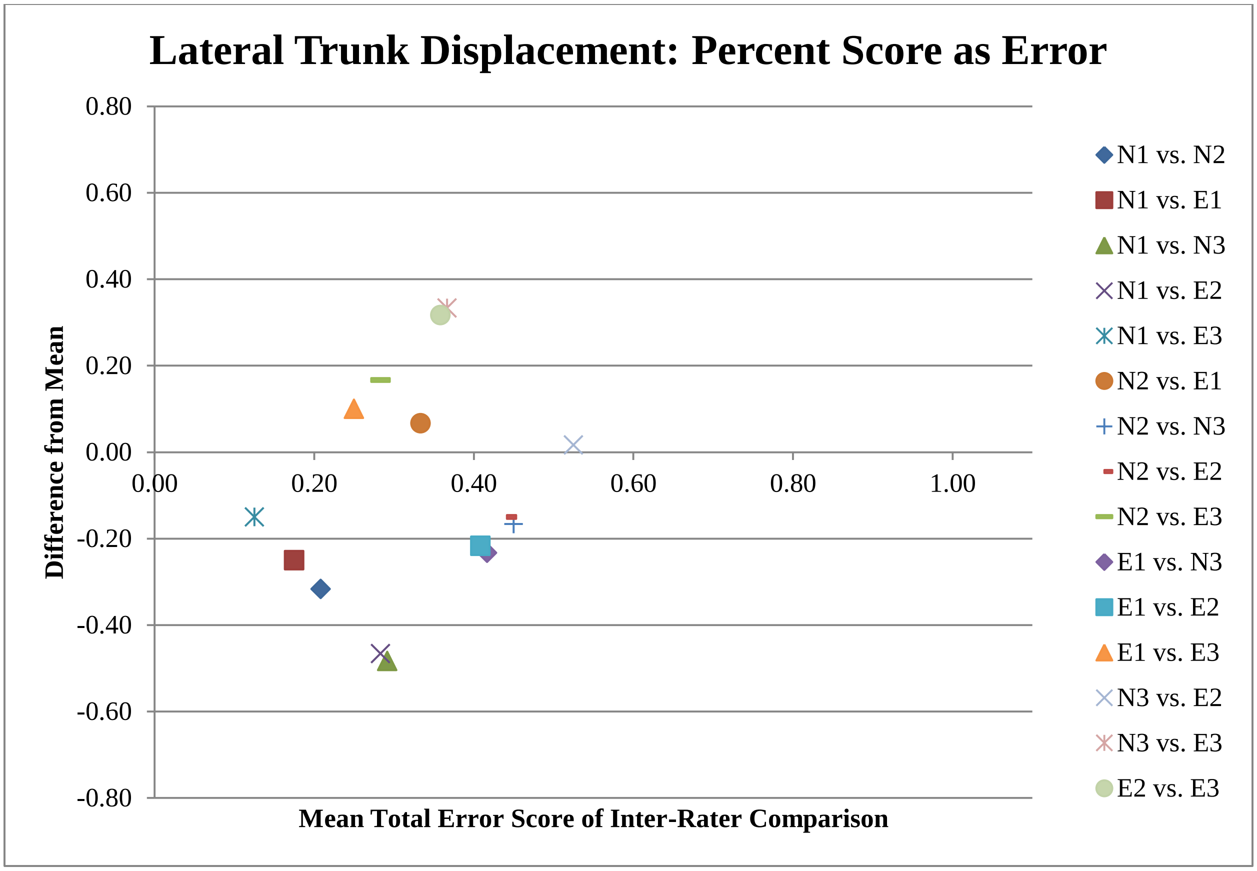

Figure 7. Bland-Altman of Lateral Trunk Displacement. Note: Overlap of data points may cause appearance of missing data points.

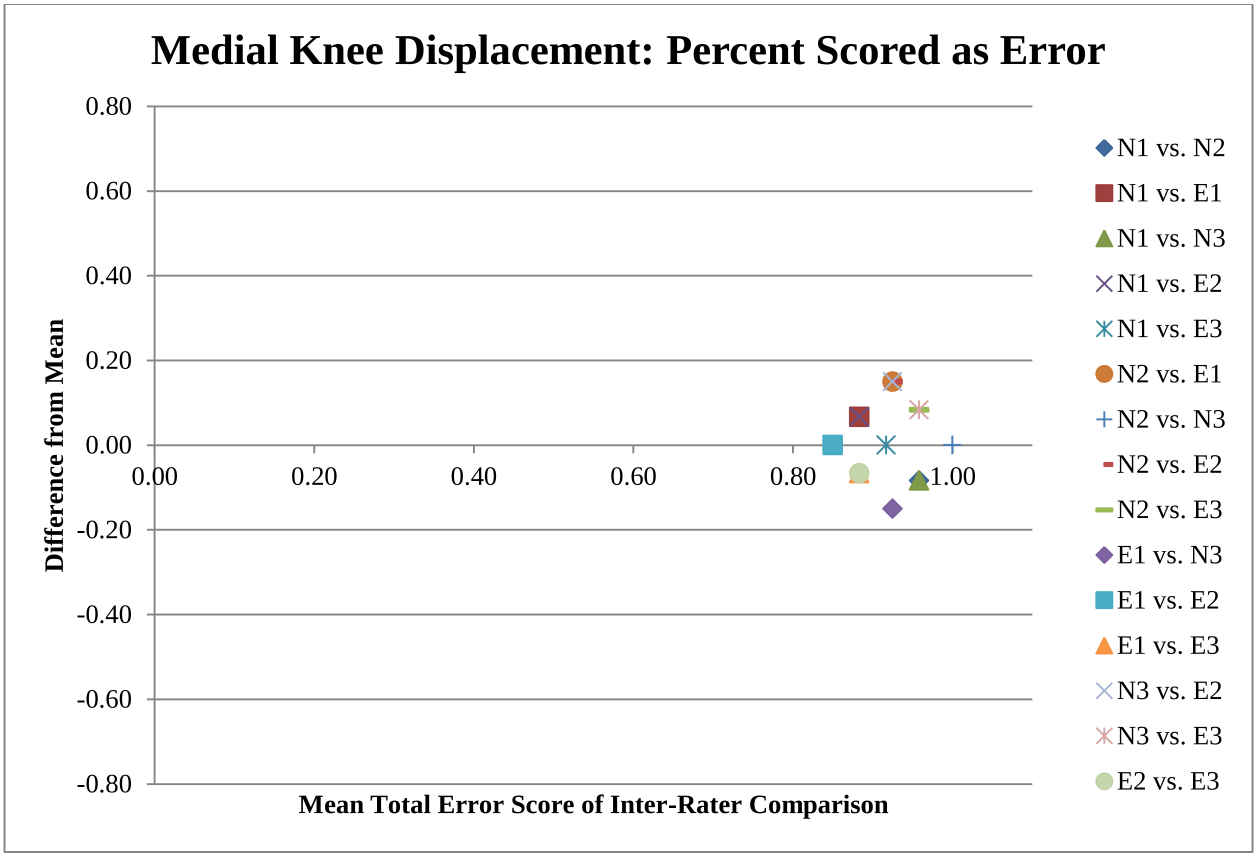

Figure 8. Bland-Altman of Medial Knee Displacement. Note: Overlap of data points may cause appearance of missing data points.

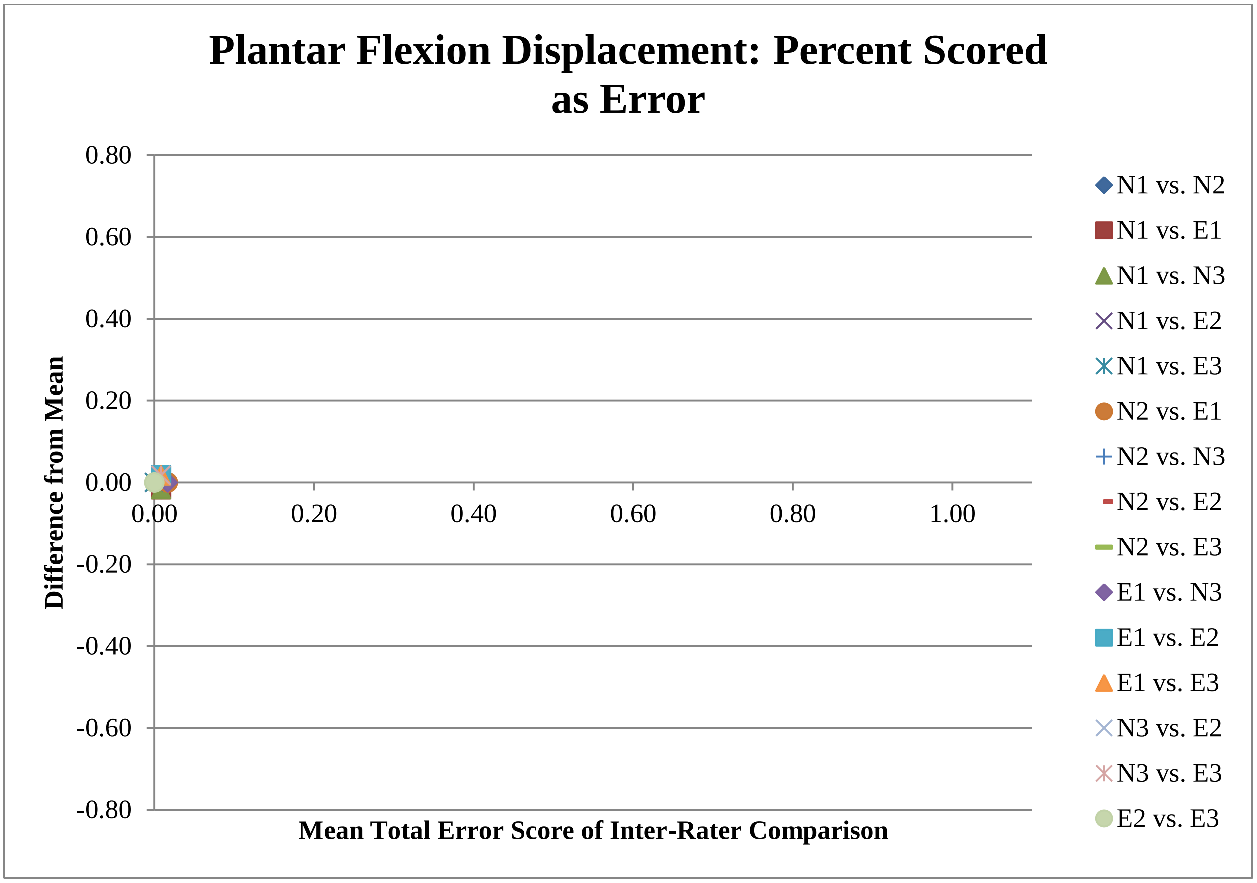

Figure 9. Bland-Altman of Plantar Flexion. Note: Overlap of data points may cause appearance of missing data points.

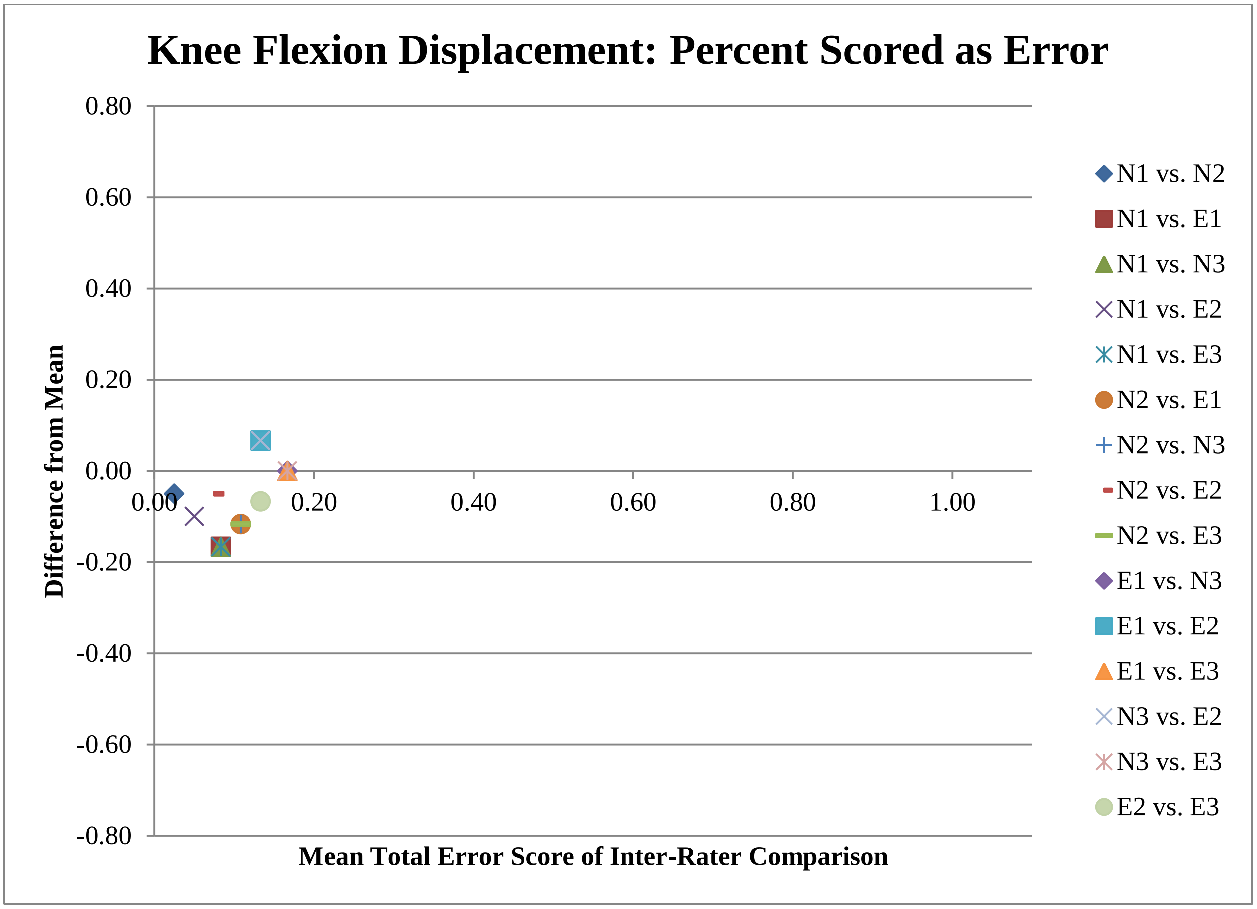

Figure 10. Bland-Altman of Knee Flexion Displacement. Note: Overlap of data points may cause appearance of missing data points.

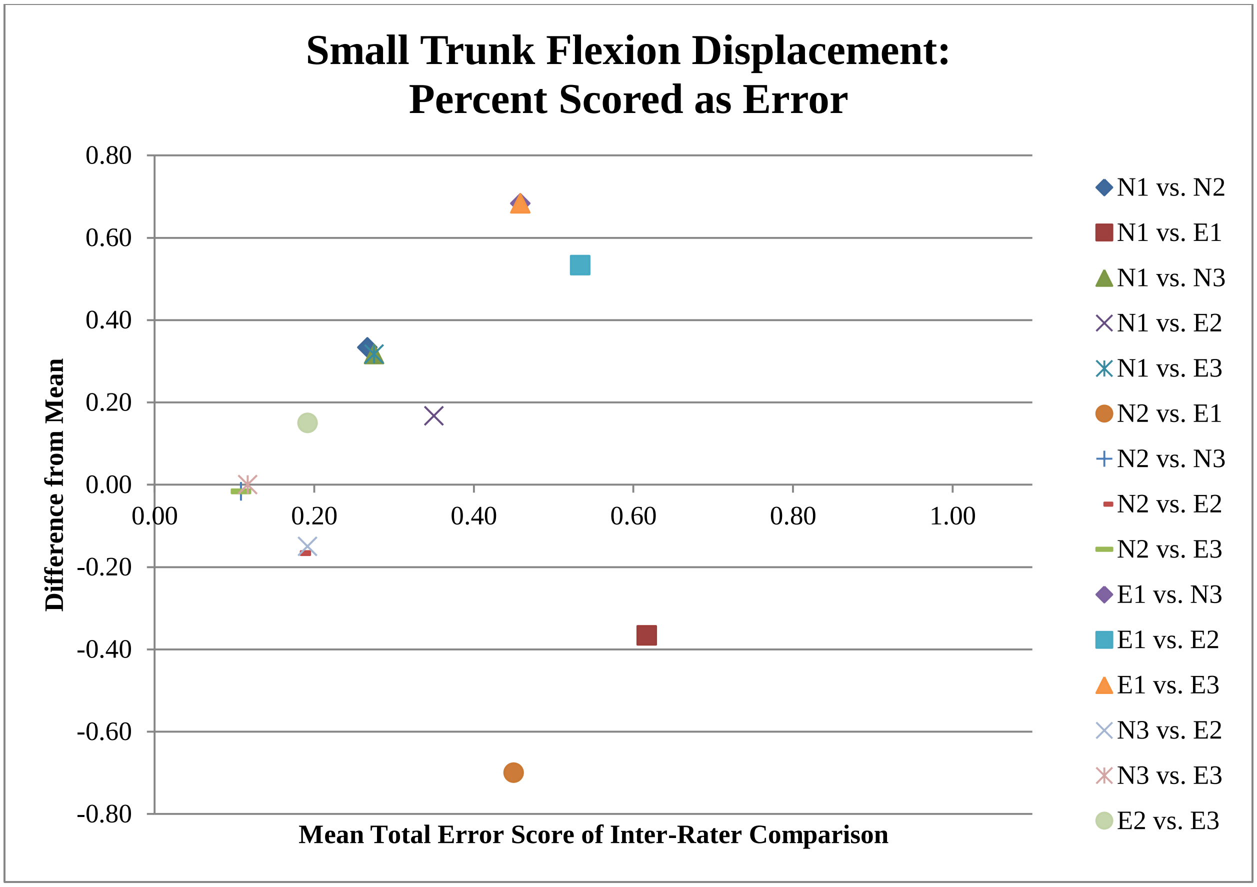

Figure 11. Bland-Altman of small trunk flexion displacement. Note: Overlap of data points may cause appearance of missing data points.

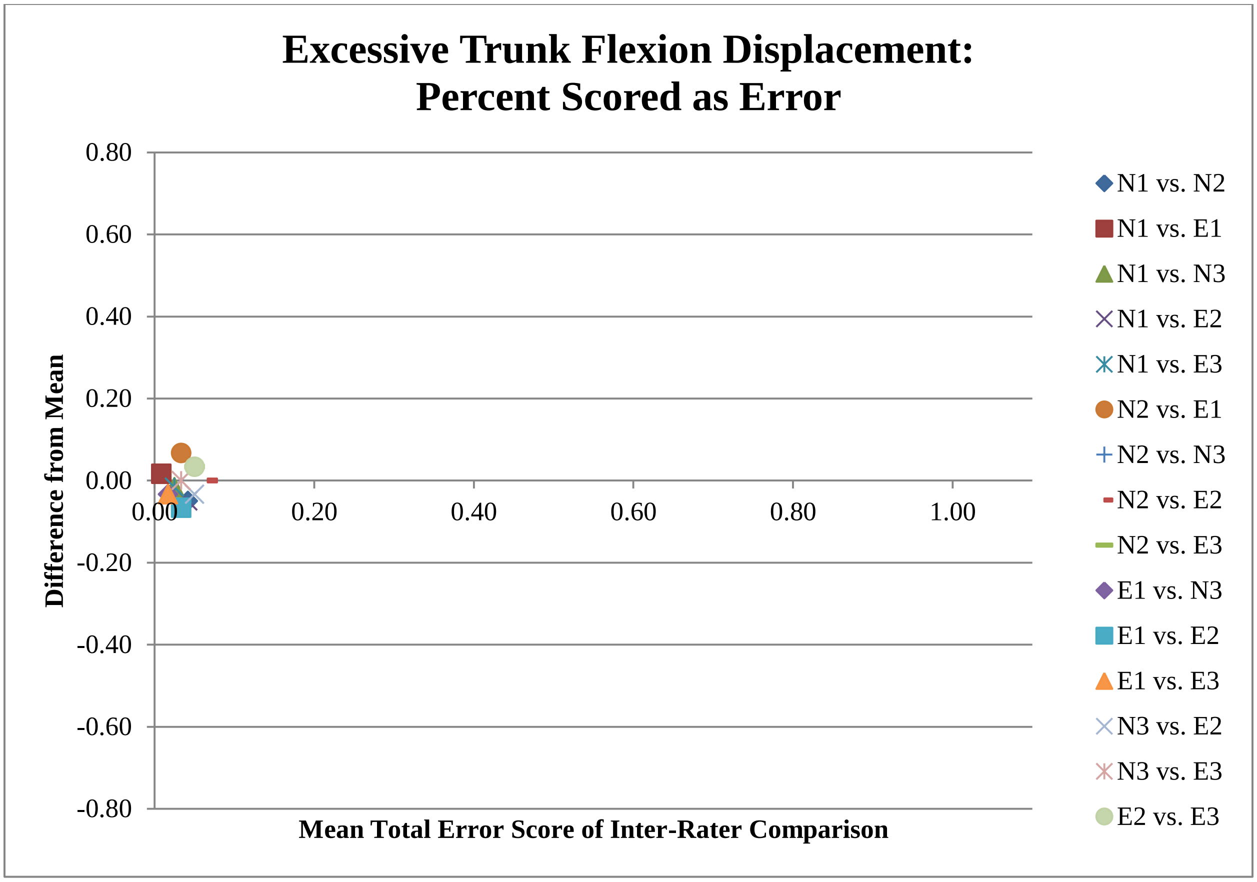

Figure 12. Bland-Altman of Excessive Trunk Flexion Displacement. Note: Overlap of data points may cause appearance of missing data points.

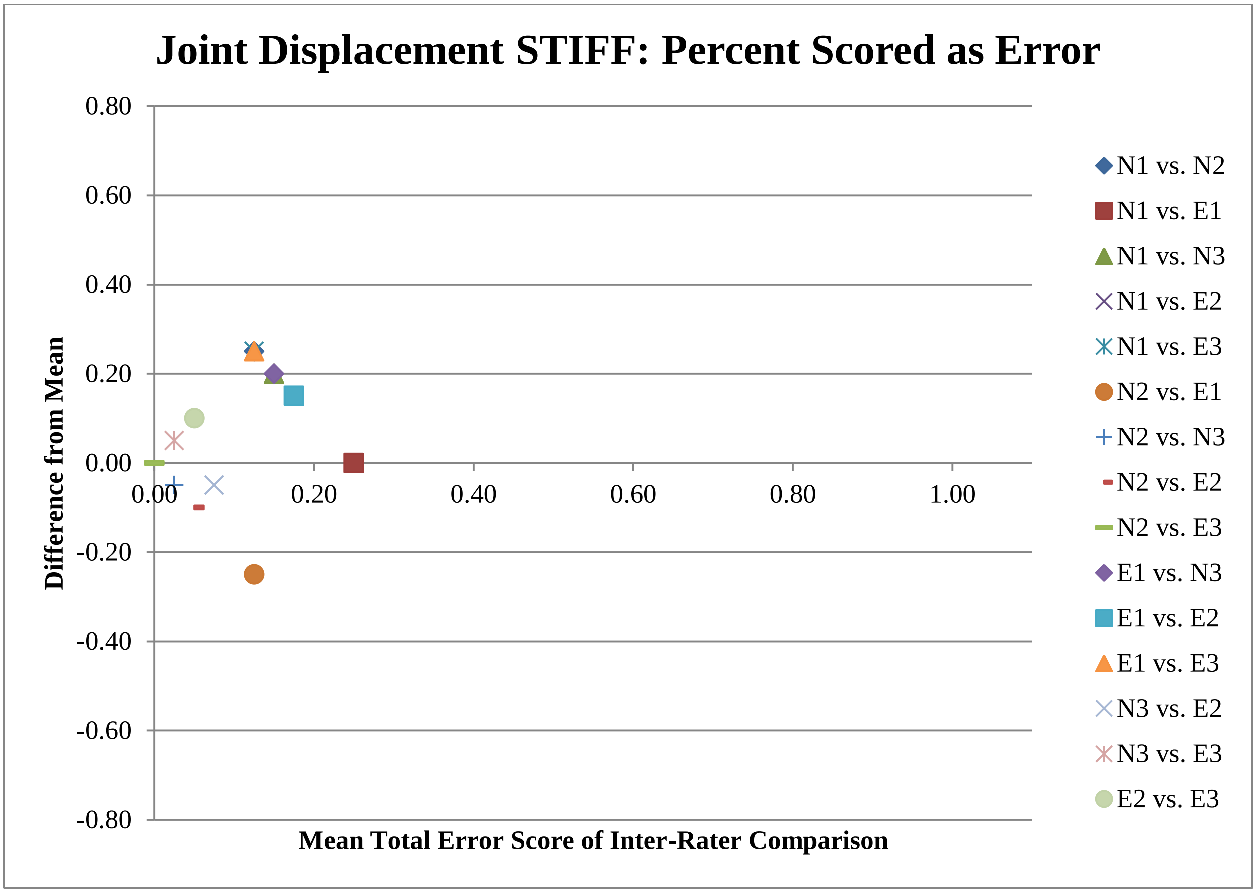

Figure 13. Bland-Altman of Joint Displacement STIFF. Note: Overlap of data points may cause appearance of missing data points.

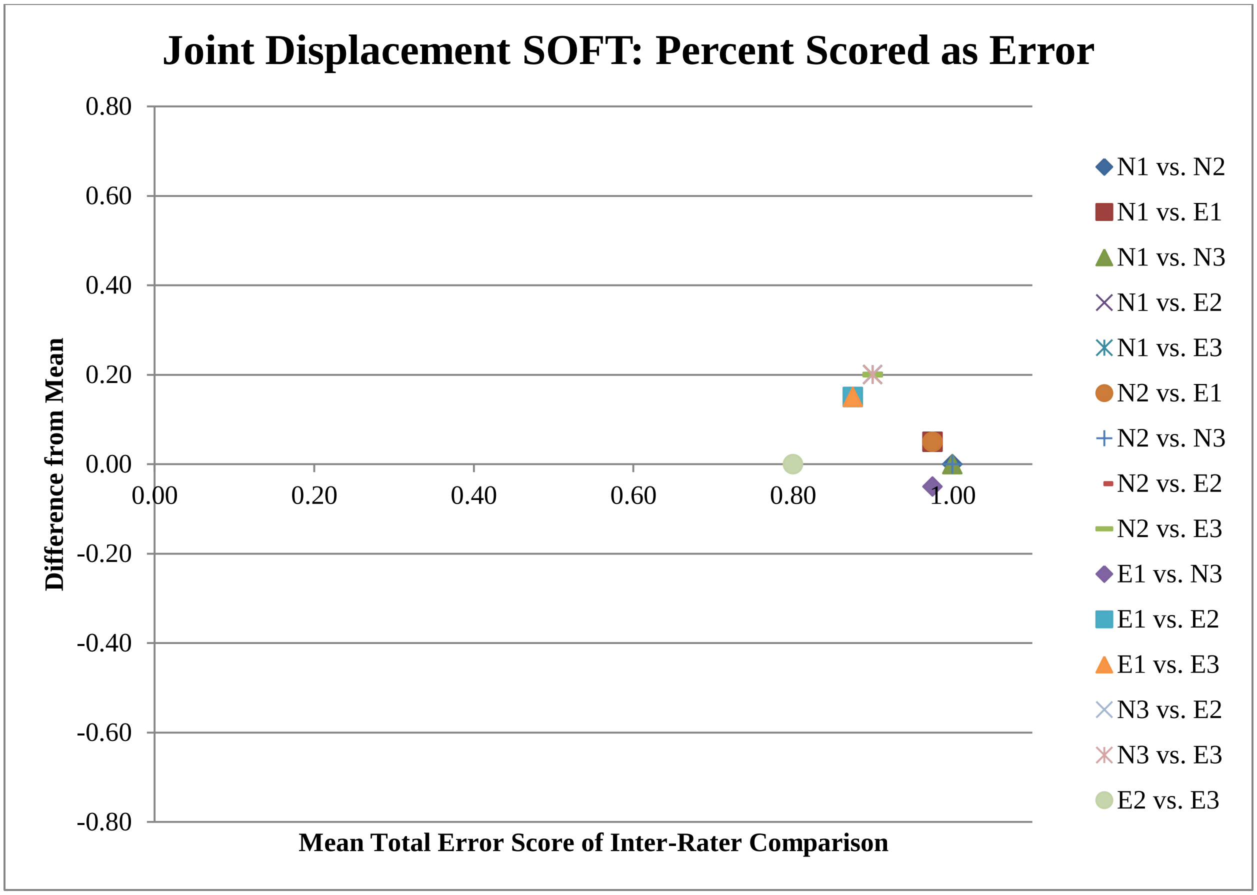

Figure 14. Bland-Altman of Joint Displacement SOFT. Note: Overlap of data points may cause appearance of missing data points.

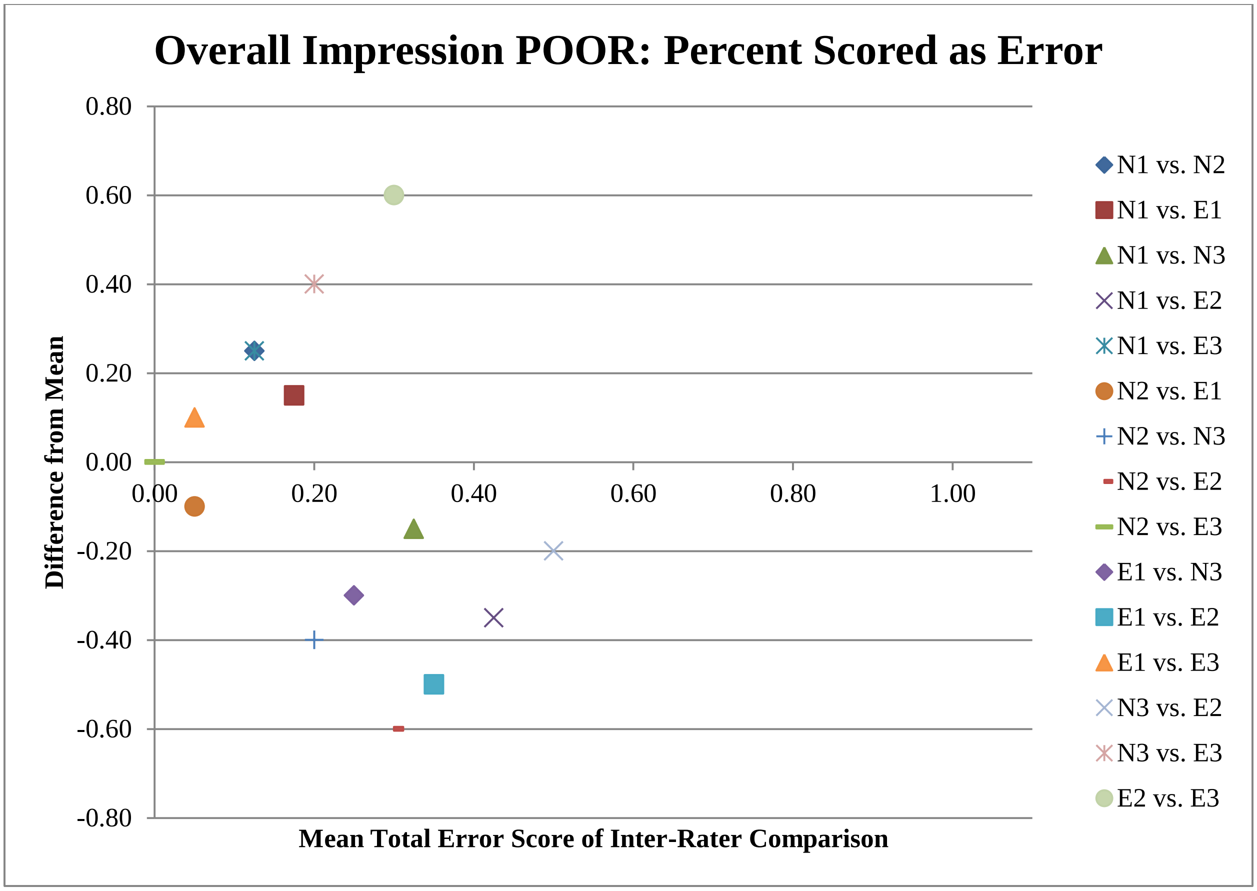

Figure 15. Bland-Altman of Overall Impression POOR. Note: Overlap of data points may cause appearance of missing data points.

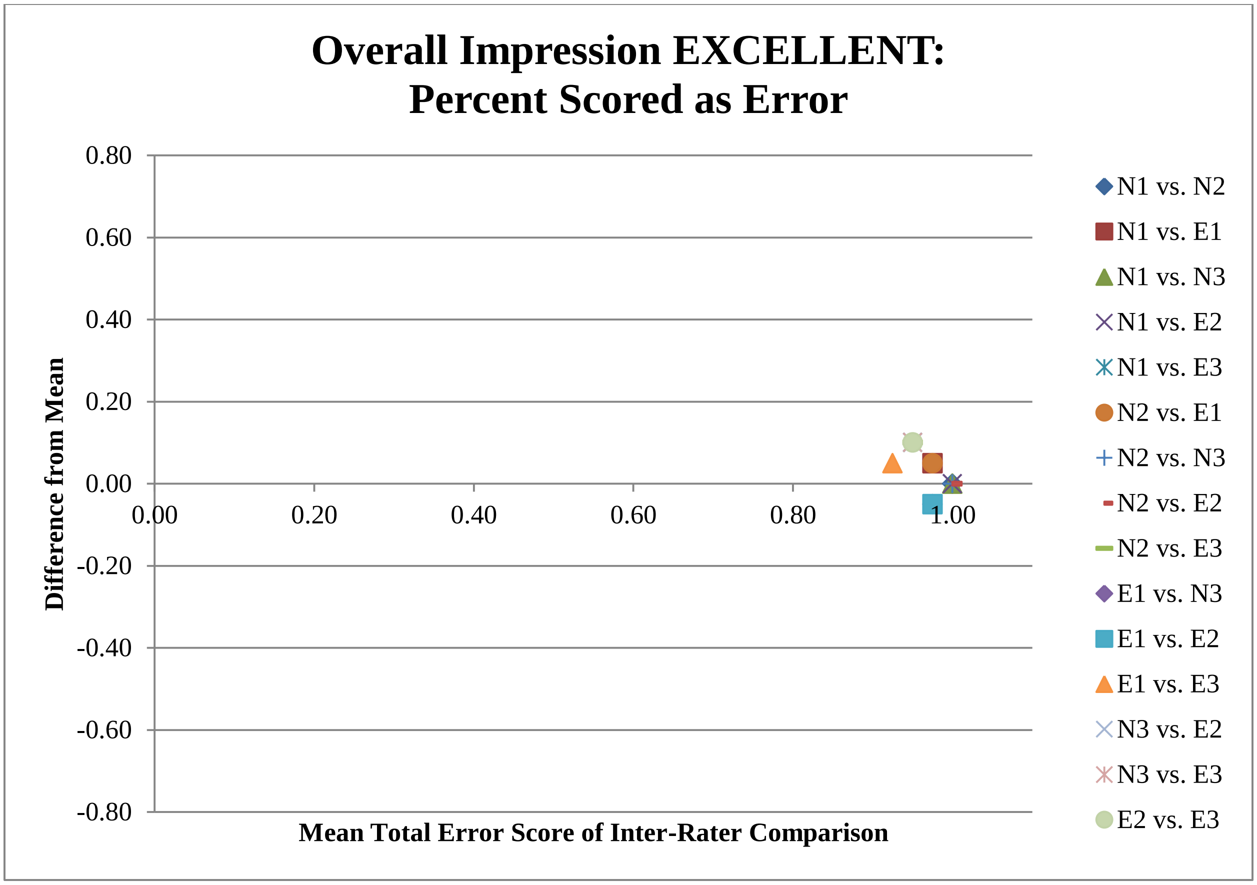

Figure 16. Bland-Altman of Overall Impression EXCELLENT. Note: Overlap of data points may cause appearance of missing data points.

Intra-rater Reliability on Average Total CST Score: Intra-class Correlation Coefficient

Table 16 presents the results of the intra-rater reliability results for the average total CST score using the ICC test for each rater. Raters ranged from having poor to excellent intra-rater reliability. The intra-rater reliability for all of the raters combined was moderate ($ICC_{(2,1)} = 0.64$, SEM = .069). The Novice raters demonstrated excellent intra-rater reliability as a group, (Novice: $ICC_{(2,1)} = 0.99$, SEM = .10) while the Veteran raters had poor intra-rater reliability (Veteran: $ICC_{(2,1)} = 0.38$, SEM = 1.72). V3 was removed from overall intra-rater comparisons amongst Novices and Veterans because this rater completed a different set of videos for the second test.

	ICC$_{(2,1)}$	SEM
N1	0.51	0.91
N2	0.90	0.31
V1	0.63	1.70
N3	0.58	0.93
V2	0.14	1.53
V3	0.95	0.18

Table 17. ICC$_{(2,1)}$ and SEM values for intra-rater reliability.

Intra-rater Reliability: Bland-Altman plots

Because V3 completed a different set of videos, two comprehensive Bland-Altman graphs were constructed comparing pre- and post-tests for all raters: one with V3 included and one with V3 removed to show the comparison of V3's reliability visually. These graphs show the intra-rater reliability for every relevant rater on each of the five videos retested. A highly correlated comprehensive graph will have the majority of data points close to 0.00 on the Y-axis. In addition, individual graphs were created comparing the pre- and post-tests of each rater for a single subject. V3 was excluded from every comparison except for video 17. Subject video 17 was the only video used by all the raters. Bland-Altman intra-rater reliability graphs are found in figures 16-22.

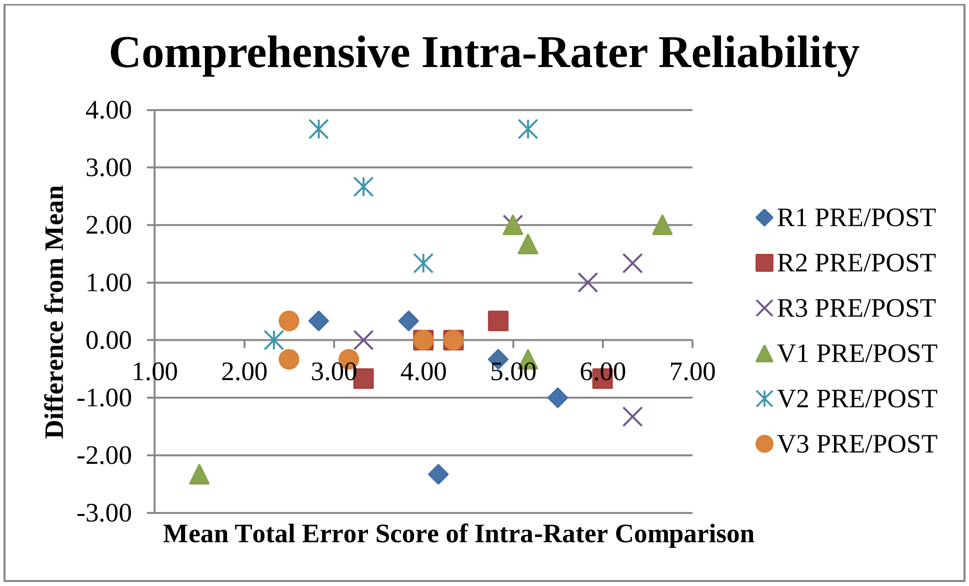

Figure 17. Bland-Altman of comprehensive Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

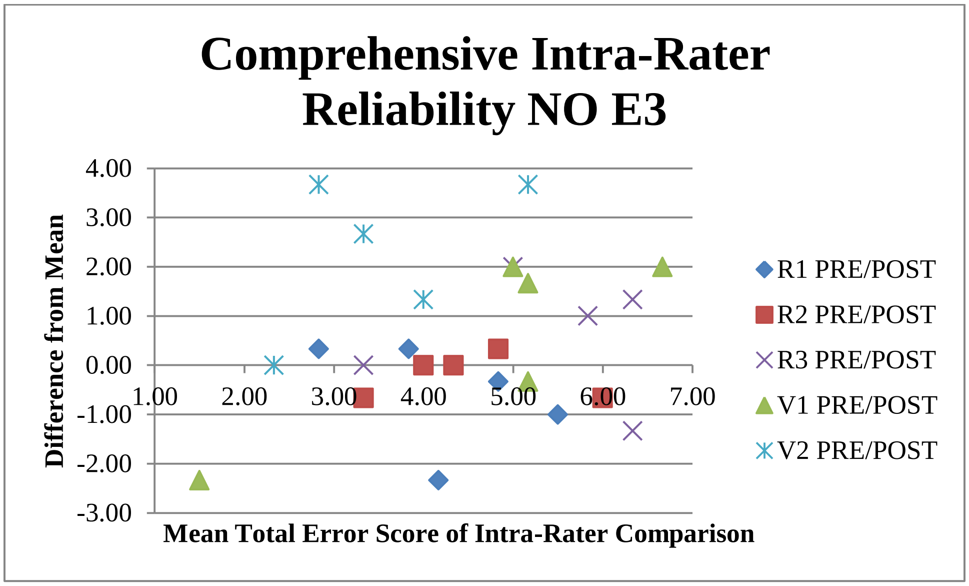

Figure 18. Bland-Altman of comprehensive Intra-Rater reliability with V3 data removed. Note: Overlap of data points may cause appearance of missing data points.

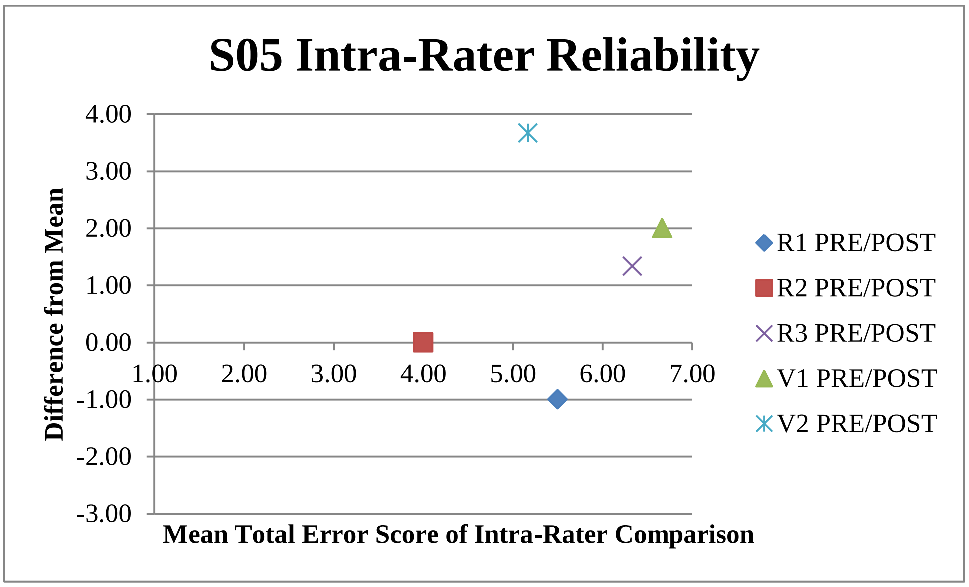

Figure 19. Bland-Altman of S05 Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

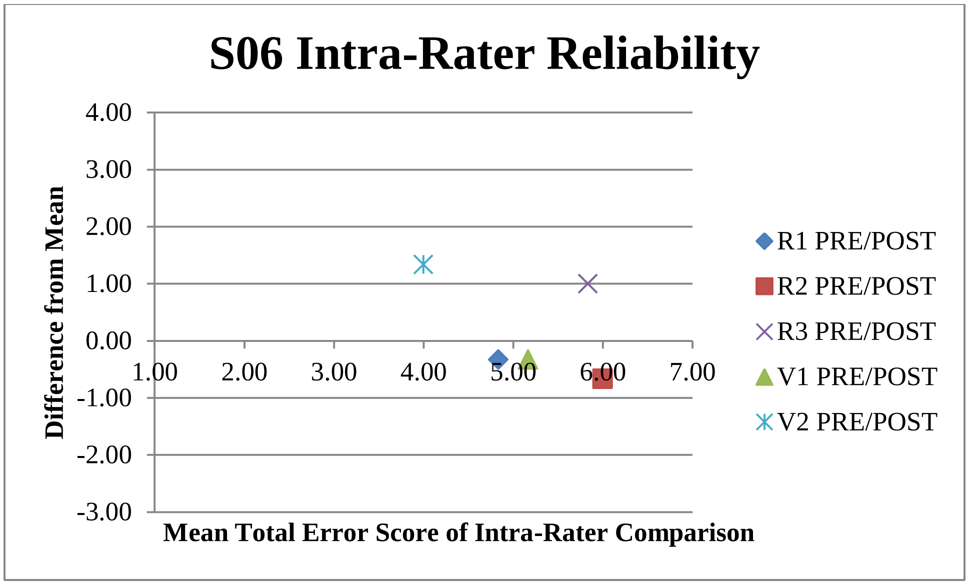

Figure 20. Bland-Altman of S06 Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

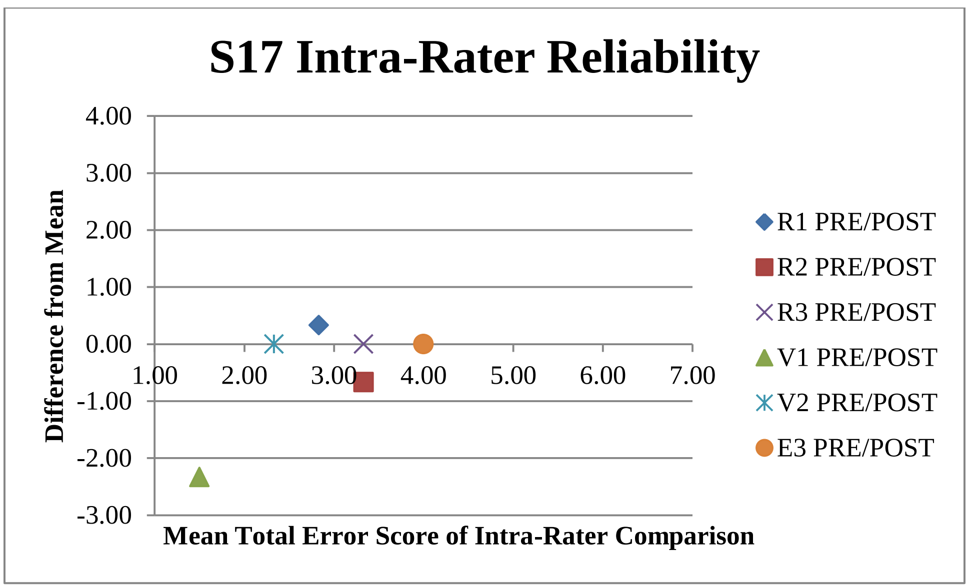

Figure 21. Bland-Altman of S17 Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

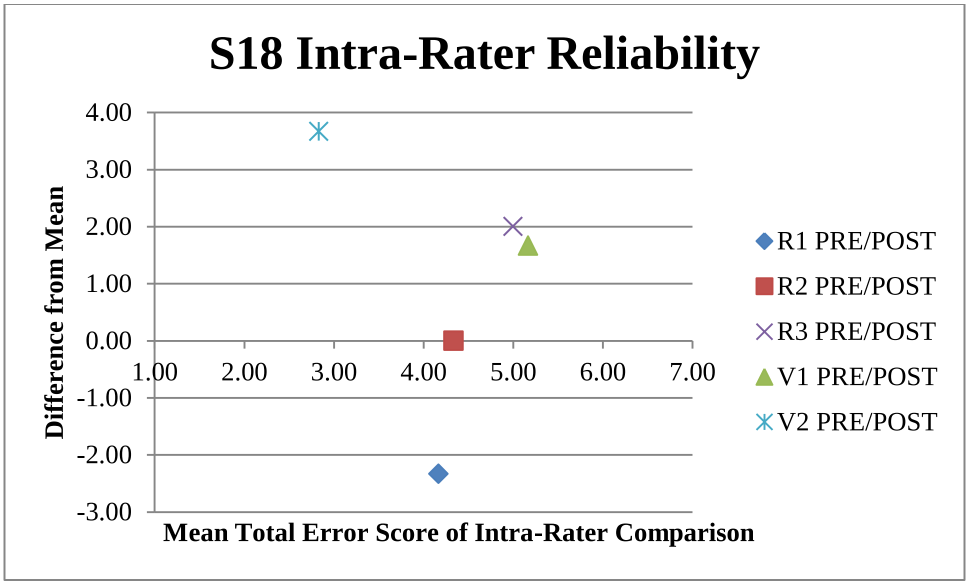

Figure 22. Bland-Altman of S18 Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

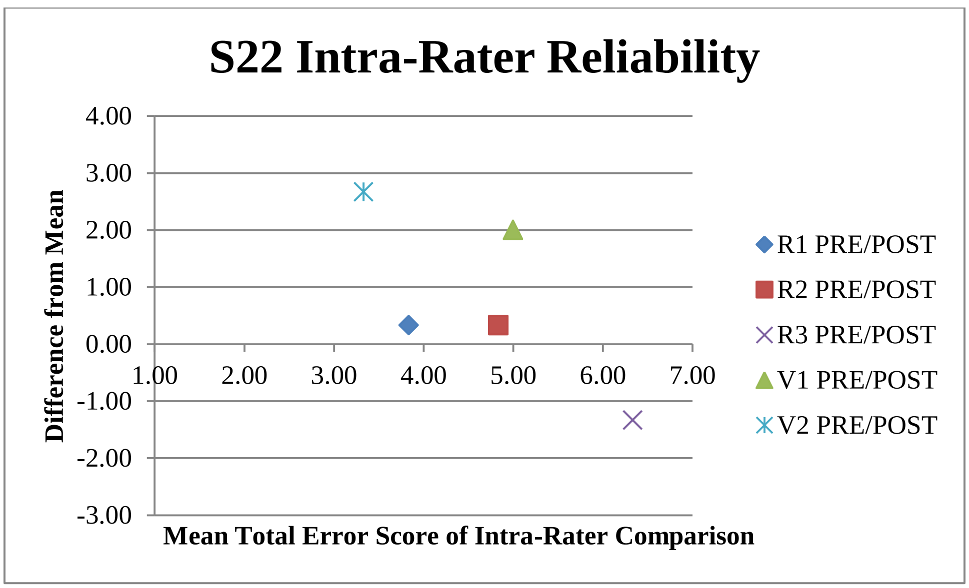

Figure 23. Bland-Altman of S22 Intra-Rater reliability. Note: Overlap of data points may cause appearance of missing data points.

Next, Bland-Altman graphs were created where the data was separated out by level of experience. Graphs were created for comprehensive data with and without V3 included. As before, graphs were also created comparing each rater for a single subject. V3 was excluded from all of these except S17 again. Novice vs. Veteran graphs are seen in Figure 23-29.

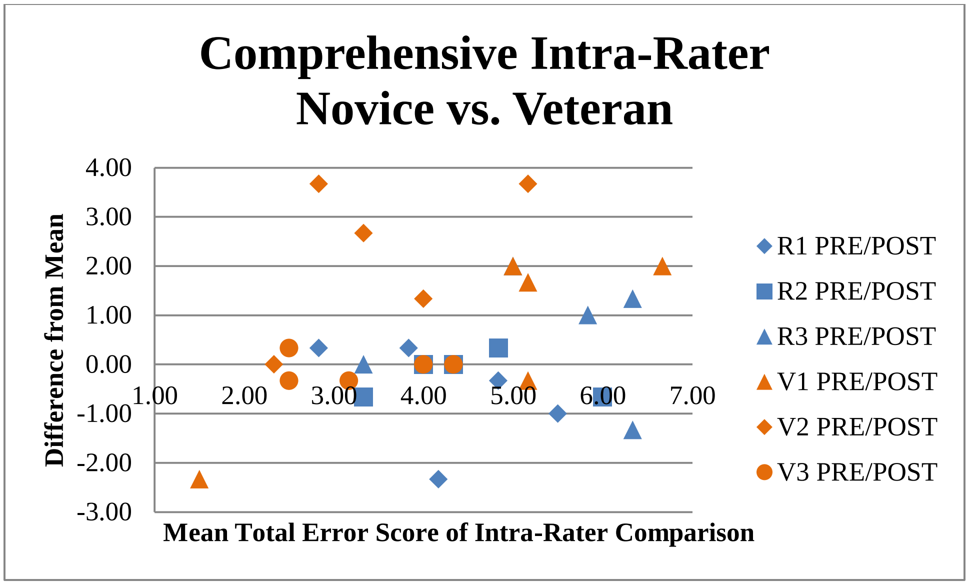

Figure 24. Bland-Altman of comprehensive Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

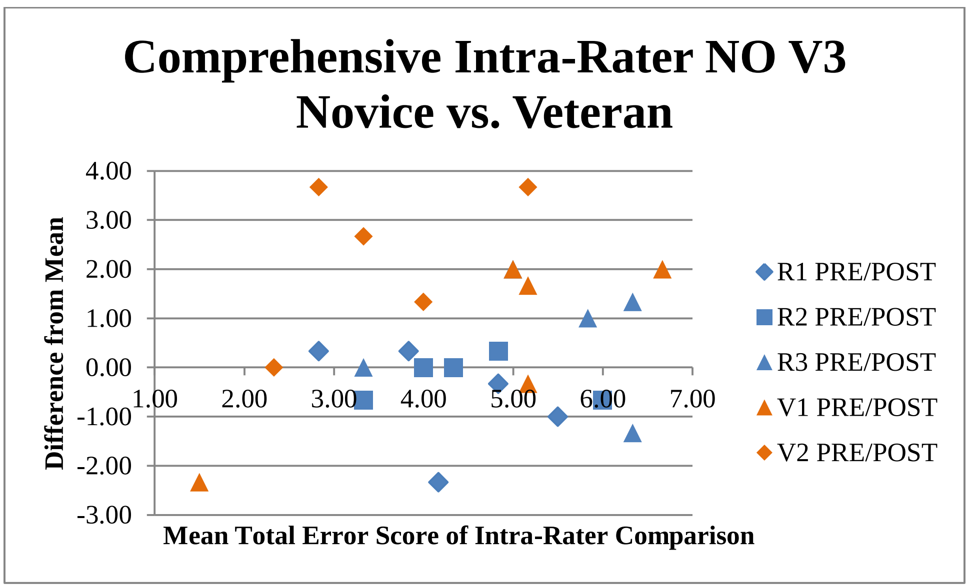

Figure 25. Bland-Altman of comprehensive Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

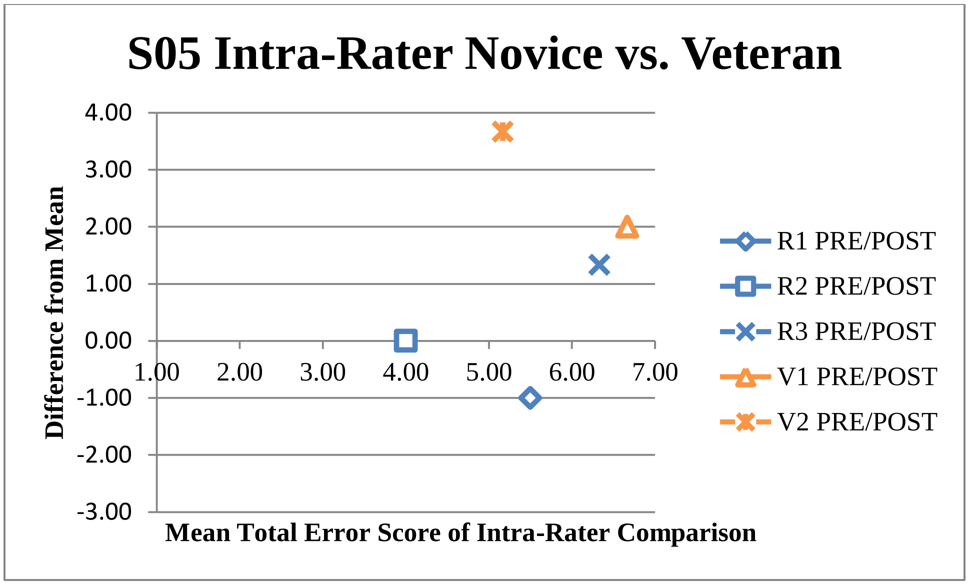

Figure 26. Bland-Altman of S05 Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

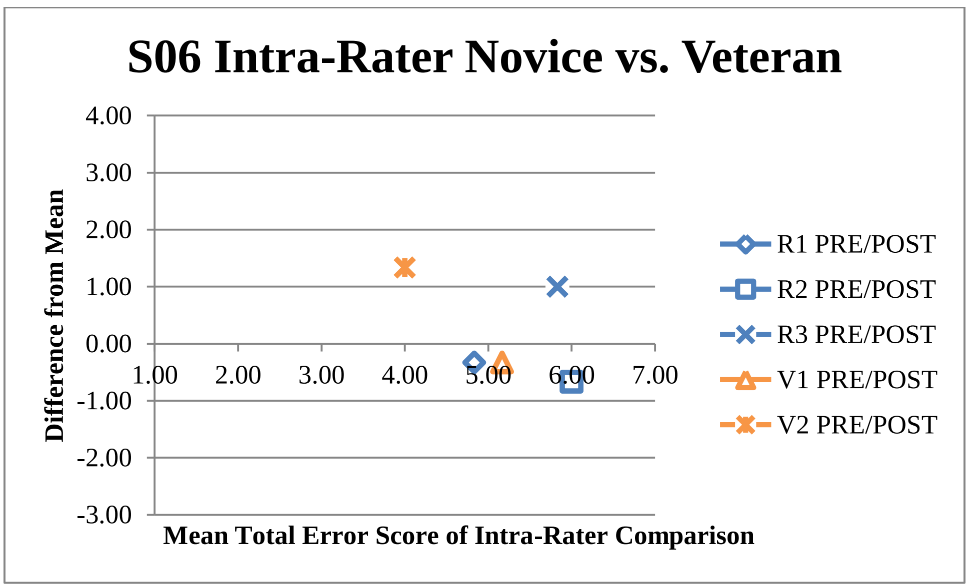

Figure 27. Bland-Altman of S06 Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

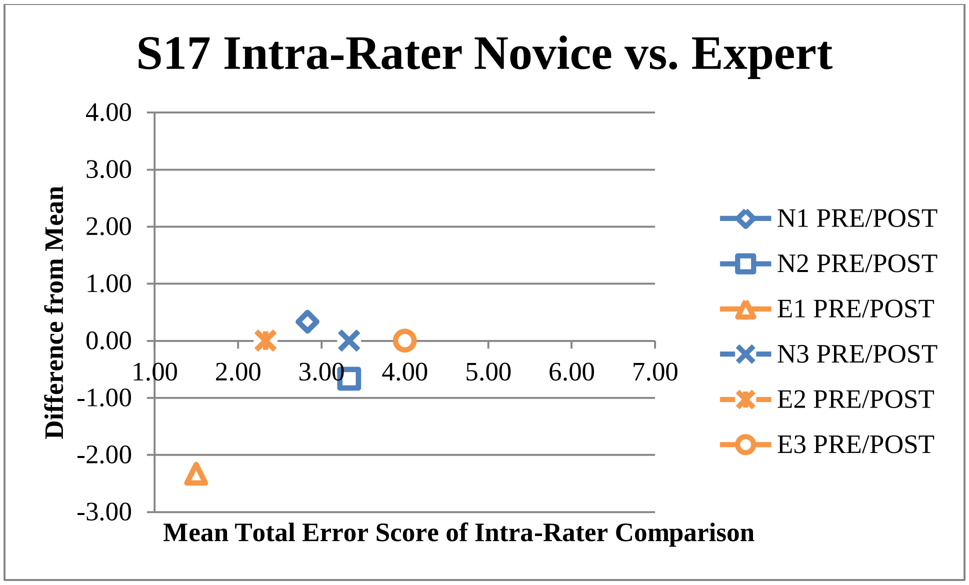

Figure 28. Bland-Altman of S17 Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

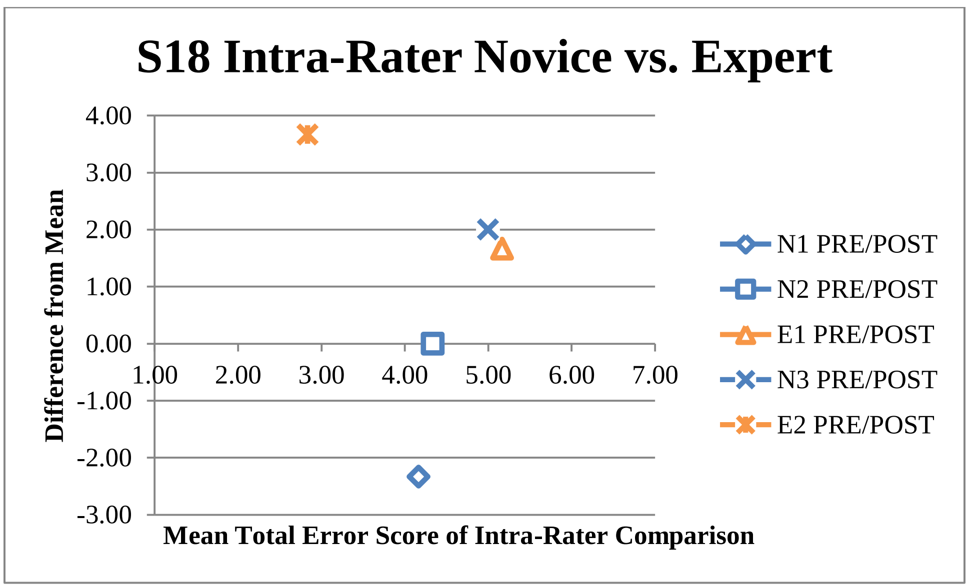

Figure 29. Bland-Altman of S18 Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

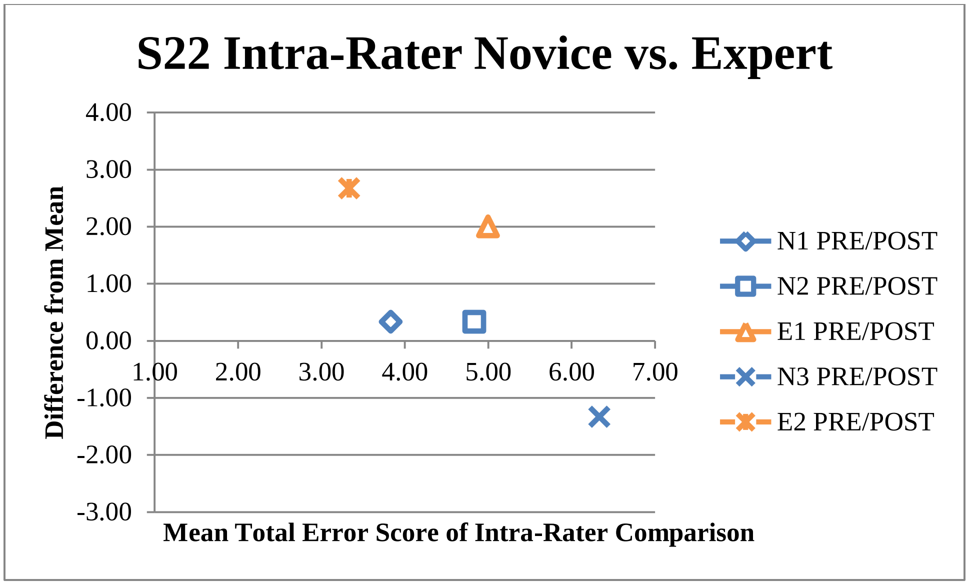

Figure 30. Bland-Altman of S22 Intra-Rater reliability separated by experience. Novice raters are blue, while Veteran raters are orange. Note: Overlap of data points may cause appearance of missing data points.

DISCUSSION AND CONCLUSIONS

The CST was designed to provide an easy, cost-efficient method to evaluate cutting biomechanics. Overall, inter-rater reliability on the CST was poor. When inter-rater reliability was calculated using only Novice vs. Veteran raters, however, the result was good. Furthermore, a majority of the individual items showed moderate-excellent inter-rater reliability, which suggests certain items on the CST may be reducing the overall reliability. Statistically, less than half of the individual items showed good-excellent inter-rater reliability; however, Cohen's kappa statistic has a flaw in its design that emerges with high chance agreement. Many of the CST's individual items had high chance agreement, negatively influencing the results. Intra-rater reliability overall was good, however, varied from excellent to poor depending on the rater's level of experience. The intra-rater reliability of the Novice group was excellent, but one rater hindered the Veteran group's reliability. The CST shows great promise as a reliable clinical screening tool if it can be further refined through clarification of training procedures and certain variables.

Current inter-rater reliability is poor. Inter-rater reliability improved when comparing via Novice vs. Veteran, which suggests that a specific rater might be the cause of the poor overall inter-rater reliability. With such a small sample, individual raters had a strong influence on the results. A larger number of participants might improve inter-rater reliability as they would give less weight to a single rater. Another method of improving inter-rater reliability would be to refine the grading process of individual items. The lowest correlation between raters was seen in N3 vs. V3, which is interesting considering they were the only two raters trained simultaneously. The highest pairwise correlations were seen between N2 vs. V2 and V1 vs. V2. Overall inter-rater reliability of the CST was lower than that of the LESS. When analyzing the LESS on the same scale that we examine the CST, the LESS had excellent inter-rater reliability compared to the poor result found in the CST. It is essential to note, however, that inclusionary criteria for participation in the CST included no formal training experience in ACL risk factor identification, which differed from previous studies involving risk factor identification.[42, 46] As a result, this study's raters potentially lacked the same insight into ACL risk factors present in other studies.

Individual item inter-rater reliability on the CST varied greatly. The most unreliable rater when compared to others on individual items was in the Veteran group (V1). The specific items showing the most variability between raters were small trunk flexion displacement, hip drop, and overall impression poor. By examining the inter-rater reliability of specific items, the overall inter-rater reliability of the CST can clearly be improved through the refinement of the weakest items. Cohen’s kappa statistic provided the most variability of data reduction analysis because the high chance agreement of individual items caused the statistic to become skewed in a negative direction. As a result, low scores were present in spite of the fact that there was an extremely high level of inter-rater agreement.

In spite of low inter-rater reliability for certain individual items when using Cohen’s kappa, there were multiple items where both kappa and the Bland-Altman agreed on reliability. Both Cohen’s kappa and the Bland-Altman statistic were in agreement that Plantar Flexion had the highest inter-rater reliability, considered excellent for both. Plantar Flexion during sidestep cutting landings has recently proven significant because it can reduce moments about the knee. [53] Previous studies looking at evaluating Plantar Flexion at initial contact with a clinical screening tool have shown excellent inter-rater reliability as well.[42, 46] According to comment and feedback sheets provided to the raters, all agreed it was the easiest item to grade.

The variable Knee Flexion Displacement also had an excellent level of agreement despite a low Cohen's kappa statistic. We believe this poor score was due to the high chance of agreement, which skewed the results negatively. The Bland-Altman plots showed an excellent correlation, supporting the hypothesis of skewed results affecting the Cohen's kappa. Our research showed similar results to those of a clinical screening tool comparing inter-rater reliability with three-dimensional motion analysis of a parallel risk factor item, showing promise for future research into the validity of the CST.[42, 46]

Multiple studies have demonstrated that a wide foot placement during a single-leg cutting motion increases stress on the ACL.[39, 41] In addition, an externally rotated foot has proven to stress the ACL more than an internally rotated foot during a single-leg cutting motion.[39] As a result, chose to count only wide foot placement and external foot rotation as errors. Foot position at initial contact had a moderate average Cohen's kappa statistic, while Bland-Altman showed good reliability compared to other items on the CST. In addition, our overall inter-rater agreement was higher than a comparison to the gold standard from a similar previous study, potentially supporting the use of the CST as a valid tool.[46] This study's results are similar to previous research and add to the body of knowledge on observing wide foot placement as previous work evaluated this variable during a double-leg task.[42, 46] Foot rotation had a moderate Cohen's kappa statistic, a comparably excellent Bland-Altman, and appropriately similar results to a study comparing inter-rater reliability to the gold standard.[42, 46]

Current research shows that poor neuromuscular control of trunk flexion is an indicator of increased ACL risk.[38] To that end, we examined two variables, small trunk flexion displacement and Excessive Trunk Flexion Displacement. Small trunk flexion displacement had a poor Cohen's kappa statistic score with a wide range of agreements. In addition, the Bland-

Altman graph showed large variability when compared to other items of the CST. This contradicts previous research involving trunk flexion displacement reliability.[42, 46] Previous research had higher reliability potentially because trunk flexion displacement was limited to one variable while the CST has two variables. By splitting trunk flexion displacement into two variables, there is increased potential for inter-rater discrepancies. Excessive trunk flexion displacement had a moderate Cohen's kappa statistic compared to small trunk flexion displacement; however this is one of the scores that was affected by the high chance agreement, as the agreement was excellent. The Bland-Altman correlation was excellent compared to other items. The excellent Bland-Altman with high agreement implies that excessive trunk flexion displacement is a reliable item of the test compared to others. There is currently no reliability research looking at excessive trunk displacement, however our excessive trunk flexion displacement results are much closer to that of previous research on trunk flexion displacement due to the high number of agreements between raters.[42, 46]

Because of the moderate correlations in Cohen's kappa, it is hard to determine whether these results will support or contradict a validity comparison with the gold standard. Overall, this study's two trunk flexion inter-rater reliability results are not as strong as previous research. Refining these variables in either operational definitions or how raters are trained to distinguish between the two could be beneficial.

Due to the necessity of binary data reduction and the potential complication of three-option classification of joint displacement and overall impression, we split these items into two categories each for data reduction. We split both into extremes (soft/average vs. stiff/average or poor/average vs. excellent/average, respectively) but left average as the comparative reference point between the data sets. During the training process, we expressed that a soft or stiff landing (for example) would be obvious, and everything else would fall into the description of “average”. Removing average as a comparison point would prevent the intended meaningful gradient between extremes when interpreting the data.

The Cohen’s kappa for joint displacement stiff and joint displacement soft were both considered poor, however the results had excellent percentage of agreement, lending credence to the idea that the kappa was again suppressed by the high chance agreement. In the Bland-Altman, joint displacement stiff had good reliability while joint displacement soft had excellent reliability. This supports the higher percentage of agreement in joint displacement soft. By looking at high inter-rater agreement, its effect on Cohen’s kappa, and the Bland-Altman correlations, we can conclude that these two items sustain the idea of good inter-rater reliability of the CST. Both of these studies support the idea that the CST will have good validity. However, corrections for the risk of high chance agreement in the kappa scores will have to be accounted for.

Overall impression of the CST also was split into two categories: overall impression poor and overall impression excellent. Overall impression poor had a poor Cohen’s kappa and a wide range of agreements. In addition, the Bland-Altman graph was poor compared to other items. It’s possible that there was confusion amongst raters as to the decision to score overall impression poor if they felt pressured to answer a specific way based on general performance on other items of the CST. Previous research had a moderate level of agreement between raters, supporting the idea that there is often rater discrepancy in how to grade a subject’s performance.[46]

Overall impression excellent had a much higher Cohen’s kappa, although still poor. However, as the lowest level of agreement in overall impression excellent was more than double that of overall impression poor, it stands to reason that the low Cohen’s kappa is a result of a high chance agreement. This is supported by the relatively excellent correlation of the Bland-Altman graph. Together both overall impression Cohen's kappa statistic would likely be similar to that of previous research in terms of agreements.[46] However, since previous research was a moderate correlation, there is still room for strengthening these items.

Previous research has prospectively shown the presence of poor lateral trunk flexion control in those who later suffered ACL or general knee ligament injuries.[38] While previous reliability studies looked at poor neuromuscular control during initial contact, we examined lateral trunk displacement, showing neuromuscular control throughout the single-leg cutting motion.[42, 46] This research showed a moderate Cohen's kappa statistic, which was less than the "good" score shown by Onate *et al.*[46] The percent of inter-rater agreement was good, but we had a wide range of agreements overall, showing that we potentially need to refine the operational definitions or testing procedures prior to further research. Another confirmed difficulty with our research occurred because our video subjects were wearing athletic wear of their own choosing; this created a difficulty in identifying the different anatomical landmarks needed to successfully identify lateral trunk displacement throughout the cutting motion. Onate *et al.*[46] showed excellent agreement with the gold standard in lateral trunk flexion at initial contact, showing room for improvement with our own research.

Hip drop was a new item created specifically for the CST. [54-56] Current research regarding strength's role in ACL injury typically involves measuring strength of the pelvic girdle with a hand-held dynamometer.[56, 57] Recent research is now showing that neuromuscular control, not strength, plays a larger role in maintaining a steady pelvic girdle.[56, 58-60] Clinical screening tools typically examine neuromuscular control of the pelvic girdle muscles with hip abduction.[42, 43, 46, 56, 61] Currently no research examines dynamic pelvic girdle neuromuscular control using another method. As a result, there is little research examining the neuromuscular control involved in maintaining a stable pelvic girdle during a cutting motion. The presence of hip drop indicates a lack of neuromuscular control of the hip abductors including the gluteus medius and would be similar to Trendelenburg's gait. Our research attempted to create another method of identifying neuromuscular control of the pelvic girdle. Cohen's kappa statistic showed moderate results, with a wide variability in inter-rater agreement. This shows that while the test itself may not be satisfactory to our needs now there is room for improvement. Multiple raters mentioned the difficulty in rating this item, as the subjects in the videos were not wearing tight clothing, making it difficult to locate the hips using typically identifiable body landmarks such as the anterior superior iliac spine and the head of the femur. Differentiating between the presence of hip internal rotation and the presence of hip drop also could improve the reliability of this item. The Bland-Altman showed poor reliability, emphasizing the need for improvement of this particular item with future research.

Medial knee displacement has shown to be a strong risk factor associated with ACL injury.[32, 34, 36, 37] Previous research shows excellent inter-rater reliability with identifying medial knee displacement in Cohen's kappa.[42, 46] Our study had poor inter-rater reliability with Cohen's kappa and Fleiss' kappa, however this item was skewed due to the high chance agreement. The average agreement was excellent amongst our raters and the Bland-Altman supports this with an excellent correlation as well. The high percentage of inter-rater agreement combined with the excellent rating from the Bland-Altman shows the CST has high inter-rater reliability on identifying medial knee displacement, a key item of the test. These results are encouraging because the purpose of this study was to show that the CST can reliably identify risk factors associated with ACL injury. By reliably identifying one of the key components of ACL injury risk, the CST has a strong foundation for future research into determining validity. These findings for inter-rater reliability are promising overall. They show that the CST may not be ready for use now, but can be improved with minimal changes.

Intra-rater reliability over all was good. When the Novice and Veteran groups were separated, however, the Novice group demonstrated excellent reliability while the Veteran group was poor. Veteran raters had a complication because of the fact that the rater with the highest overall intra-rater reliability (V3) was removed from a majority of the equations due to an error in grading videos. This individual used a different set of videos from the rest of the raters, rendering this rater's data invalid for the group intra-rater comparisons. Had this rater been included in the intra-rater reliability for Veteran raters, the reliability may have improved.

Novice raters had excellent intra-rater reliability, which was significant. We hypothesize the Novice raters were more reliable overall than the Veteran raters because they had more a recent formal education, affecting the way they look at injuries. Veteran raters were certified at the time when internship programs still existed, potentially limiting the amount of formal in-class education they received. In addition, an emphasis on quality of movement in research and likely education has only recently occurred. As a result, Veteran raters were more likely to rely on the anecdotal evidence they had picked up through the years. This is supported by the fact that V1 was the most inconsistent with other raters on scoring individual items. In contrast to this, however, the most consistent rater was Veteran as well, which could imply it is not the CST as a video screening tool but rather the raters whom are unreliable. Another consideration is the fact that there were only two males participating, both Veterans. Future research could look at the difference between Novice and Veteran when the groups are balanced across genders.

One difficulty facing our raters was the general lack of formal movement screening experience compared to past studies looking at the reliability of clinical screening tools. Onate *et al.*[46] used only two raters, one with 15 years' experience and one with approximately one year experience when testing inter-rater reliability of the LESS. However, the Veteran rater was part of the research team that developed the LESS, and the Veteran rater trained the Novice rater, which gave their study an advantage as to what was desired. Onate, *et al.*[46] did not test intra-rater reliability. Padua, *et al.*[42]only used two raters testing the LESS as well, however, there is little to nothing mentioned about their level of experience as an athletic trainer, with ACL injuries, or with movement screening. Most likely, both raters were Veteran athletic trainers involved in the development of the LESS. Our study did not have the same level of experience with ACL research and formal movement screening that the other studies did, as we were specifically attempting to teach athletic trainers with no experience to use the CST to reliably identify ACL risk factors.

We designed our test with an eye to imitating the reliability research done by previous ACL injury prevention studies Padua *et al.* (reference) and Onate *et al.* (reference) as well as expanding them with a larger sample size and complete reliability testing. By limiting our study to purely intra- and inter-rater reliability, we hoped to expand on the idea of reliability as an important part of tests identifying ACL risk factors.

Padua *et al.*[42] required a minimum of one week between testing sessions for intra-rater reliability, while we only required a minimum of three days for the CST. Compared to this previous study, there was a greater risk of learning retention overlap when grading the second set of videos compared to the first set in our research.

FUTURE RESEARCH

Future research of the CST should evaluate the validity of the CST compared with three-dimensional motion analysis. Increasing the number of raters to see if a higher number affects the reliability could also be effective. Altering the demographics of the raters to include more males across multiple experience levels, changing the number of years experienced as an athletic trainer, or potentially including athletic trainers with a history of ACL prevention research or training in movement screening would all be beneficial. Performing validity and reliability testing on different potential items to either include or remove from the CST could help increase the effectiveness of the CST at identifying risk factors. A limitation of this study is that while we evaluated many more raters than previous research with movement screening, we still only had a small number of raters to truly evaluate the role of experience in using the CST. Conducting the study again with a larger number of subjects for intra-rater reliability would be beneficial. Also, as previously mentioned, subjects filmed for the videos were wearing standard athletic wear of their choosing, leading to a variety of visibility of anatomical body parts. Filming subjects wearing tighter-fitting clothing with anatomical landmarks visible and/or identified could possibly lead to improved refining, definition, or recreation of the CST components.

In conclusion, the CST has a strong foundation for identifying ACL risk factors at this time. While it still needs refining, the key elements for inter- and intra-rater reliability are present. Medial knee displacement, foot rotation, plantar flexion, knee flexion displacement, and joint displacement are all reliable areas within the test. Increasing the reliability of the other items would help solidify the CST's effectiveness. The CST will become a powerful clinical screening tool for use in preventing ACL injuries.

REFERENCES:

1. Gianotti SM, Marshall SW, Hume PA, Bunt L. Incidence of anterior cruciate ligament injury and other knee ligament injuries: A national population-based study. *J Sci Med Sport*. 2009:622--627.

2. De LoÃ«s M, Dahlstedt LJ, ThomÃ©e R. A 7-year study on risks and costs of knee injuries in male and female youth participants iN12 sports. *Scandinavian Journal of Medicine and Science in Sports*. 2000;10:90-97.

3. Marshall SW, Padua DA, McGrath M. Incidence of ACL injury. In: *Understanding and Preventing Non-Contact ACL Injury.* Champaign, IL: Human Kinetics; 2007.

4. Hootman JM, Dick R, Agel J. Epidemiology of Collegiate Injuries 15 Sports: Summary and Recommendations for Injury Prevention Initiatives. *J Athl Train*. 2007;42:311-312-319.

5. Silvers HJ, Mandelbaum BR. Prevention of anterior cruciate ligament injury in the female athlete. *Br J Sports Med*. 2007;41 Suppl 1:i52-9.

6. Nelson F, Billinghurst RC, Pidoux I, et al. Early post-traumatic osteoarthritis-like changes in human articular cartilage following rupture of the anterior cruciate ligament. *Osteoarthritis Cartilage*. 2006;14:114-119.

7. Myer GD, Ford KR, Hewett TE. Rationale and clinical techniques for anterior cruciate ligament injury prevention among female athletes. *Journal of Athletic Training*. 2004;39:352-364.

8. Cumps E, Verhagen E, Annemans L, Meeusen R. Injury rate and socioeconomic costs resulting from sports injuries in Flanders: data derived from sports insurance statistics 2003. *Br J Sports Med*. 2008;42:767-772.

9. Manske RC, Prohaska D, Lucas B. Recent advances following anterior cruciate ligament reconstruction: rehabilitation perspectives : Critical reviews in rehabilitation medicine. *Curr Rev Musculoskelet Med*. 2012.

10. Castanharo R, da Luz BS, Bitar AC, D'Elia CO, Castropil W, Duarte M. Males still have limb asymmetries in multijoint movement tasks more thaN2 years following anterior cruciate ligament reconstruction. *J Orthop Sci*. 2011;16:531-535.

11. Mattacola CG, Perrin DH, Gansneder BM, Gieck JH, Saliba EN, McCue FC,3rd. Strength, Functional Outcome, and Postural Stability After Anterior Cruciate Ligament Reconstruction. *J Athl Train*. 2002;37:262-268.

12. Neitzel JA, Kernozek TW, Davies GJ. Loading response following anterior cruciate ligament reconstruction during the parallel squat exercise. *Clin Biomech (Bristol, Avon)*. 2002;17:551-554.

13. Colby SM, Hintermeister RA, Torry MR, Steadman JR. Lower limb stability with ACL impairment. *J Orthop Sports Phys Ther*. 1999;29:444-51; discussion 452-4.

14. Paterno MV, Ford KR, Myer GD, Heyl R, Hewett TE. Limb asymmetries in landing and jumping 2 years following anterior cruciate ligament reconstruction. *Clin J Sport Med*. 2007;17:258-262.

15. Paterno MV, Schmitt LC, Ford KR, Rauh MJ, Myer GD, Hewett TE. Effects of sex on compensatory landing strategies upon return to sport after anterior cruciate ligament reconstruction. *J Orthop Sports Phys Ther*. 2011;41:553-559.

16. Brown TD, Johnston RC, Saltzman CL, Marsh JL, Buckwalter JA. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. *J Orthop Trauma*. 2006;20:739-744.

17. Lohmander LS, Ostenberg A, Englund M, Roos H. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. *Arthritis Rheum*. 2004;50:3145-3152.

18. Kessler MA, Behrend H, Henz S, Stutz G, Rukavina A, Kuster MS. Function, osteoarthritis and activity after ACL-rupture: 11 years follow-up results of conservative versus reconstructive treatment. *Knee Surg Sports Traumatol Arthrosc*. 2008;16:442-448.

19. Lopiano DA. Modern history of women in sports - Twenty-five years of Title IX. *Clinics in sports medicine*. 2000;19:163-173.

20. Haycock CE, Gillette JV. Susceptibility of women athletes to injury. Myths vs reality. *JAMA*. 1976;236:163-165.

21. Gomez E, DeLee JC, Farney WC. Incidence of injury in Texas girls' high school basketball. *Am J Sports Med*. 1996;24:684-687.

22. Agel J, Olson DE, Dick R, Arendt EA, Marshall SW, Sikka RS. Descriptive epidemiology of collegiate women's basketball injuries: National Collegiate Athletic

Association injury surveillance system, 1988-1989 through 2003-2004. *Journal of Athletic Training*. 2007;42:202-210.

23. Loud KJ, Micheli LJ. Common athletic injuries in adolescent girls. *Curr Opin Pediatr*. 2001;13:317-322.

24. Damore DT, Metzl JD, Ramundo M, Pan S, Van Amerongen R. Patterns in childhood sports injury. *Pediatr Emerg Care*. 2003;19:65-67.

25. Shea KG, Pfeiffer R, Jo HW, Curtin M, Apel PJ. Anterior cruciate ligament injury in pediatric and adolescent soccer players: An analysis of insurance data. *Journal of Pediatric Orthopaedics*. 2004;24:623-628.

26. Boden BP, Torg JS, Knowles SB, Hewett TE. Video analysis of anterior cruciate ligament injury: Abnormalities in hip and ankle kinematics. *Am J Sports Med*. 2009;37:252-259.

27. Pantano KJ, White SC, Gilchrist LA, Leddy J. Differences in peak knee valgus angles between individuals with high and low Q-angles during a single limb squat. *Clin Biomech*. 2005;20:966-972.

28. Chandrashekar N, Slauterbeck J, Hashemi J. Sex-based differences in the anthropometric characteristics of the anterior cruciate ligament and its relation to intercondylar notch geometry: a cadaveric study. *Am J Sports Med*. 2005;33:1492-1498.

29. Chandrashekar N, Mansouri H, Slauterbeck J, Hashemi J. Sex-based differences in the tensile properties of the human anterior cruciate ligament. *J Biomech*. 2006;39:2943-2950.

30. Arendt EA, Bershadsky B, Agel J. Periodicity of noncontact anterior cruciate ligament injuries during the menstrual cycle. *J Gend Specif Med*. 2002;5:19-26.

31. Slauterbeck JR, Fuzie SF, Smith MP, et al. The Menstrual Cycle, Sex Hormones, and Anterior Cruciate Ligament Injury. *J Athl Train*. 2002;37:275-278.

32. Markolf KL, Burchfield DM, Shapiro MM, Shepard MF, Finerman GAM, Slauterbeck JL. Combined knee loading states that generate high anterior cruciate ligament forces. *J Orthop Res*. 1995;13:930-935.

33. Ireland ML. Anterior cruciate ligament injury in female athletes: Epidemiology. *Journal of Athletic Training*. 1999;34:150-154.

34. Bere T, Florenes TW, Krosshaug T, et al. Mechanisms of anterior cruciate ligament injury in World Cup alpine skiing: a systematic video analysis of 20 cases. *Am J Sports Med*. 2011;39:1421-1429.

35. Krosshaug T, Slauterbeck JR, Engebretsen L, Bahr R. Biomechanical analysis of anterior cruciate ligament injury mechanisms: three-dimensional reconstruction from video sequences. *Scand J Med Sci Sports*. 2007;17:508--519.

36. Krosshaug T, Nakamae A, Boden BP, et al. Mechanisms of anterior cruciate ligament injury in basketball: video analysis of 39 cases. *Am J Sports Med*. 2007;35:359-367.

37. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med*. 2005;33:492-501.

38. Zazulak BT, Hewett TE, Reeves NP, Goldberg B, Cholewicki J. Deficits in neuromuscular control of the trunk predict knee injury risk: a prospective biomechanical-epidemiologic study. *Am J Sports Med*. 2007;35:1123-1130.

39. Dempsey AR, Lloyd DG, Elliott BC, Steele JR, Munro BJ, Russo KA. The effect of technique change on knee loads during sidestep cutting. *Med Sci Sports Exerc*. 2007;39:1765-1773.

40. Borotikar BS, Newcomer R, Koppes R, McLean SG. Combined effects of fatigue and decision making on female lower limb landing postures: central and peripheral contributions to ACL injury risk. *Clin Biomech (Bristol, Avon)*. 2008;23:81-92.

41. Dempsey AR, Lloyd DG, Elliott BC, Steele JR, Munro BJ. Changing sidestep cutting technique reduces knee valgus loading. *Am J Sports Med*. 2009;37:2194-2200.

42. Padua DA, Marshall SW, Boling MC, Thigpen CA, Garrett WE,Jr, Beutler AI. The Landing Error Scoring System (LESS) Is a valid and reliable clinical assessment tool of jump-landing biomechanics: The JUMP-ACL study. *Am J Sports Med*. 2009;37:1996-2002.

43. Barber-Westin SD, Smith ST, Campbell T, Noyes FR. The drop-jump video screening test: retention of improvement in neuromuscular control in female volleyball players. *J Strength Cond Res*. 2010;24:3055-3062.

44. Ekegren CL, Miller WC, Celebrini RG, Eng JJ, Macintyre DL. Reliability and validity of observational risk screening in evaluating dynamic knee valgus. *J Orthop Sports Phys Ther*. 2009;39:665-674.

45. Ford KR, Myer GD, Hewett TE. Reliability of landing 3D motion analysis: Implications for longitudinal analyses. *Med Sci Sports Exerc*. 2007;39:2021-2028.

46. Onate J, Cortes N, Welch C, Van Lunen BL. Expert versus novice interrater reliability and criterion validity of the landing error scoring system. *J Sport Rehabil*. 2010;19:41-56.

47. Distefano LJ, Padua DA, Distefano MJ, Marshall SW. Influence of age, sex, technique, and exercise program on movement patterns after an anterior cruciate ligament injury prevention program in youth soccer players. *Am J Sports Med*. 2009;37:495-505.

48. Conn JM, Annest JL, Gilchrist J. Sports and recreation related injury episodes in the US population, 1997-99. *Inj Prev*. 2003;9:117-123.

49. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. *Am J Sports Med*. 2007;35:1756-1769.

50. Onate JA, Guskiewicz KM, Marshall SW, Giuliani C, Yu B, Garrett WE. Instruction of jump-landing technique using videotape feedback: altering lower extremity motion patterns. *Am J Sports Med*. 2005;33:831-842.

51. DiStefano LJ, Padua DA, Blackburn JT, Garrett WE, Guskiewicz KM, Marshall SW. Integrated Injury Prevention Program Improves Balance and Vertical Jump Height in Children. *Journal of Strength and Conditioning Research*. 2010;24:332-342.

52. McLean SG, Walker K, Ford KR, Myer GD, Hewett TE, van den Bogert AJ. Evaluation of a two dimensional analysis method as a screening and evaluation tool for anterior cruciate ligament injury. *Br J Sports Med*. 2005;39:355-362.

53. Cortes N, Morrison S, Van Lunen BL, Onate JA. Landing technique affects knee loading and position during athletic tasks. *J Sci Med Sport*. 2012;15:175-181.

54. Cochrane JL, Lloyd DG, Besier TF, Elliott BC, Doyle TLA, Ackland TR. Training affects knee kinematics and kinetics in cutting maneuvers in sport. *Med Sci Sports Exerc*. 2010;42:1535-1544.

55. Bencke J, Naesborg H, Simonsen EB, Klausen K. Motor pattern of the knee joint muscles during side-step cutting in European team handball. Influence on muscular co-ordination after an intervention study. *Scand J Med Sci Sports*. 2000;10:68-77.

56. Beutler AI, de la Motte SJ, Marshall SW, Padua DA, Boden BP. Muscle strength and qualitative jump-landing differences in male and female military cadets: The jump-ACL study. *Journal of Sports Science and Medicine*. 2009;8:663-671.

57. Kollock Jr. RO, Onate JA, Van Lunen B. The reliability of portable fixed dynamometry during hip and knee strength assessments. *Journal of Athletic Training*. 2010;45:349-356.

58. Shultz SJ, Nguyen A-, Leonard MD, Schmitz RJ. Thigh strength and activation as predictors of knee biomechanics during a drop jump task. *Med Sci Sports Exerc*. 2009;41:857-866.

59. Bennett DR, Blackburn JT, Boling MC, McGrath M, Walusz H, Padua DA. The relationship between anterior tibial shear force during a jump landing task and quadriceps and hamstring strength. *Clin Biomech*. 2008;23:1165-1171.

60. Herman DC, Weinhold PS, Guskiewicz KM, Garrett WE, Yu B, Padua DA. The effects of strength training on the lower extremity biomechanics of female recreational athletes during a stop-jump task. *Am J Sports Med*. 2008;36:733-740.

61. Noyes FR, Barber-Westin S, Fleckenstein C, Walsh C, West J. The drop-jump screening test: Difference in lower limb control by gender and effect of neuromuscular training in female athletes. *Am J Sports Med*. 2005;33:197-207.

Appendix A

Video:_____________ Graded By:_________

Cutting Screening Tool: CST 1

Note: Please review the CST Instruction Sheet prior to scoring individuals with the CST.

ITEM	SCORE			CAMERA VIEW
1. Foot Placement: *Foot is lateral to greater trochanter*	☐ No Error (0)	☐ Error (1)		Front
2. Foot Rotation: *Foot is externally rotated more thaN30 degrees*	☐ No Error (0)	☐ Error (1)		Front
3. Hip Drop: *Drops from IC to MKF*	☐ No Error (0)	☐ Error (1)		Front
4. Lateral Trunk Displacement: *Trunk flexes away from the direction of the cut*	☐ No Error (0)	☐ Error (1)		Front
5. Medial Knee Displacement: *Knee medial to midfoot*	☐ No Error (0)	☐ Error (1)		Front
6. Plantar Flexion at Initial Contact: *Land heel to toe (or) flat foot*	☐ No Error (0)	☐ Error (1)		Side
7. Knee Flexion Displacement: *Knee flexes less than 45 degrees*	☐ No Error (0)	☐ Error (1)		Side
8. Small Trunk Flexion Displacement: *Trunk DOES NOT flex more than at initial contact*	☐ No Error (0)	☐ Error (1)		Side
9. Excessive Trunk Flexion Displacement: *Trunk flexes past parallel with the lower leg*	☐ No Error (0)	☐ Error (1)		Side
10. Joint Displacement: *Sagittal Plane*	☐ Soft (0)	☐ Average (1)	☐ Stiff (2)	Side
11. Overall Impression:	☐ Excellent (0)	☐ Average (1)	☐ Poor (2)	Front, Side